

L Number	Hits	Search Text	DB	Time stamp
-	29449	image with (encrypt\$4 encod\$4)	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 16:57
-	19176	image with (encrypt\$4 encod\$4) and process	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 16:57
-	24393	image with (encrypt\$4 encod\$4) and process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 16:58
-	19514	image with (encrypt\$4 encod\$4) and process\$4 with image	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:33
-	14094	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:35
-	4471	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:36
-	1750	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size))	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:40
-	0	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 adj mean	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:41
-	82	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:42
-	56	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:43
-	51	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/24 17:45
-	1	707/\$.ccsl. and image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 08:29
-	51	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fujitsu.as.	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 08:24

-	51	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as.	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:09
-	3	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and first adj process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 08:38
-	0	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and first adj process\$4 and second adj process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 08:39
-	3	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and (multiple various plurality) adj process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 08:54
-	31	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and (multiple various plurality) with process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 08:57
-	19	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and (multiple various plurality) near4 process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:00
-	16	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) near4 image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and (multiple various plurality) near4 process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:05
-	1	image with (encrypt\$4 encod\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and (multiple various plurality) adj process	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:03

-	0	image with encrypt\$4 and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) near4 image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and (multiple various plurality) near4 process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:07
-	0	image with encrypt\$4 and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 slic\$4 break\$4) near4 image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and (multiple various plurality) near4 process\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:07
-	1	image with encrypt\$4 and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 slic\$4 break\$4) near4 image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as.	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:08
-	2	image with encrypt\$4 and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) with image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4, and position with image and image with information and siz\$4 not fuji.as.	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:15
-	6	image with (encrypt\$4 scrambl\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) near4 image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as.	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 09:57
-	2	image with (encrypt\$4 scrambl\$4) and process\$4 with image and (input\$4 read\$4) with image and (divid\$4 break\$4) near4 image and (boundar\$4 (divid\$4 near4 size)) and mark\$4 with identif\$4 and position with image and image with information and siz\$4 not fuji.as. and divid\$4 with pre\$determin\$4	USPAT; US-PGPUB; EPO; JPO; IBM_TDB	2003/07/25 10:00



[11] **Patent Number:** **5,502,576**

[45] **Date of Patent:** **Mar. 26, 1996**

102(b)

- [54] **METHOD AND APPARATUS FOR THE TRANSMISSION, STORAGE, AND RETRIEVAL OF DOCUMENTS IN AN ELECTRONIC DOMAIN**
- [75] **Inventors:** **Thomas E. Ramsay**, Minneapolis;  
**James C. Elkins**, Golden Valley, both  
of Minn.
- [73] **Assignee:** **Ramsay International Corporation**,  
Minneapolis, Minn.
- [21] **Appl. No.:** **933,623**
- [22] **Filed:** **Aug. 24, 1992**
- [51] **Int. Cl.<sup>6</sup>** ..... **H04N 1/40; H04N 1/00**
- [52] **U.S. Cl.** ..... **358/444; 358/403; 358/404**
- [58] **Field of Search** ..... **358/401, 403,**  
**358/404, 443, 444, 462, 467; 345/2**

Excerpt from *MacWorld* Magazine, Aug. 1992, 3 pages.  
Excerpt from *New Media* Magazine, May, 1992, pp. 12–16.  
Excerpt from *ABA Journal*, Sep. 1992, 2 pages.  
Technology 2001, The Future of Computing and Communications, Leebaert, Derek; Chapters 7–9.

(List continued on next page.)

*Primary Examiner*—Scott A. Rogers  
*Assistant Examiner*—Jerome Grant  
*Attorney, Agent, or Firm*—Briggs and Morgan

[57] **ABSTRACT**

A method and apparatus for high speed conversion of tangible source documents to electronic images, and subsequent transmission or storage and retrieval of images, utilizing hybrid signal processing. The system employs a higher bandwidth analog signal for image capture and lower effective bandwidth analog signal for transmission or storage and retrieval, with an intervening digital memory utilized to construct a bitmap of the image to facilitate various dissection and seaming schemes which optimize image content and processing time. The system is designed around a conventional bus structure, and the memory serves as a junction with conventional personal computers, networks, and peripheral devices. In a representative embodiment, a tangible image is captured using a camera producing an analog signal with conventional raster synchronization. The synchronization is stripped from the analog signal, which is digitized for 8-bit grayscale and multiplexed to the memory where the image exists as a bitmap that may be divided into segments. The content is read from the memory, converted to an analog signal, and control signals are added. The control signals include horizontal and vertical sync pulses and interval blanking, a pilot signal to maintain alignment between adjacent segments along seams and to compensate for time-based errors, and calibration pulses to permit instantaneous adjustment of the gray level for each line, ensure accurate image content, and permit display enhancement. The resultant analog signal is stored on a randomly accessible storage medium as one or more frames, transmitted and reassembled, or displayed on a conventional monitor.

## U.S. PATENT DOCUMENTS

2,072,527	3/1937	Nicolson .....	369/275.1
2,540,105	2/1951	Dunbar et al. ....	369/275.1
2,588,680	3/1952	Williams .....	369/18
3,810,174	5/1974	Heard et al. ....	342/185
3,964,064	6/1976	Brandao et al. ....	342/185

(List continued on next page.)

## FOREIGN PATENT DOCUMENTS

0007655	7/1907	European Pat. Off. ....	G11B	7/00
0122094	3/1984	European Pat. Off. ....	H04N	5/76
WO90/04837	10/1989	WIPO .....	G08K	9/30

## OTHER PUBLICATIONS

Four page brochure from Folsom Research on Video/Scan and CGCI.

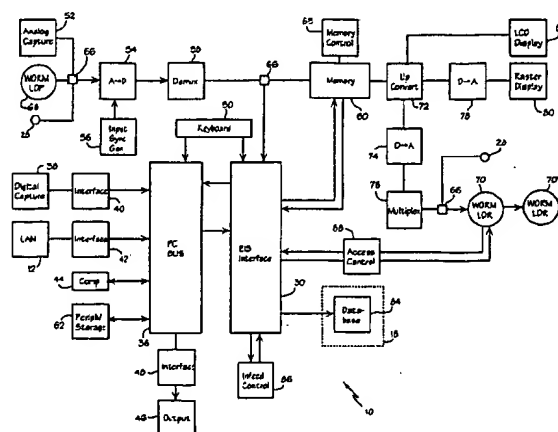
Fourteen pages of brochures from RGB Spectrum.

Six page brochure from RGB Spectrum entitled *Videographics Report*.

Excerpt from *New Media Magazine*, Aug. 1992, pp. 1, 14-17.

Excerpt from *New Media Magazine*, Sep. 1992, pp. 1, 24-27.

**51 Claims, 6 Drawing Sheets**



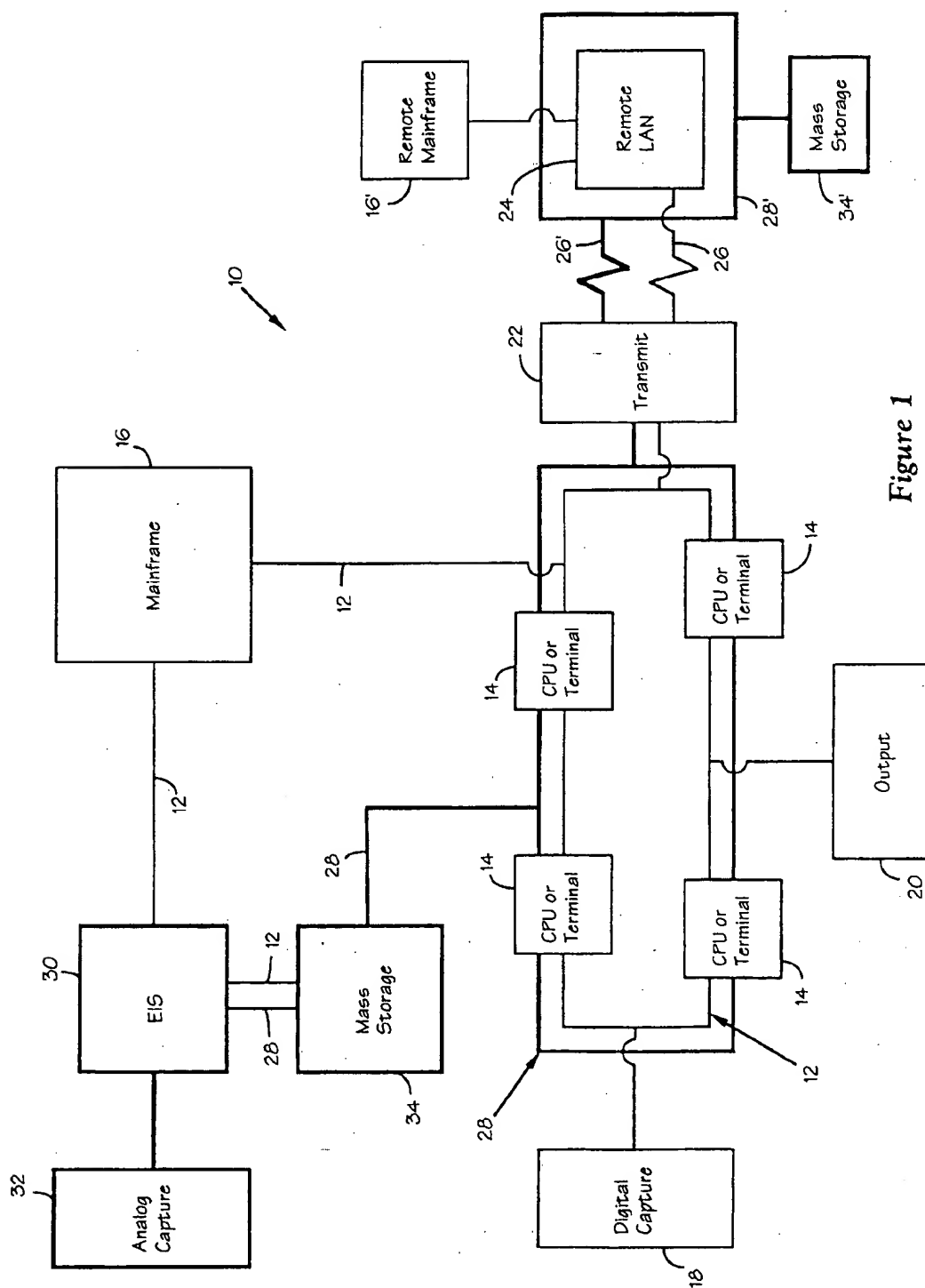
## U.S. PATENT DOCUMENTS

3,999,008	12/1976	Bouwhuis et al.	358/128
4,023,185	5/1977	Bloom et al.	369/275.1
4,097,895	6/1978	Spong	369/275.1
4,118,734	10/1978	Bouwhuis et al.	369/111
4,128,838	12/1978	Brands et al.	342/185
4,161,752	7/1979	Basilico	358/128
4,172,386	10/1979	Cribbs et al.	73/618
4,204,433	5/1980	Cribbs et al.	73/618
4,232,376	11/1980	Dion et al.	365/222
4,236,221	11/1980	Cribbs et al.	348/163
4,275,425	6/1981	Engle	360/92
4,313,188	1/1982	Bartolini et al.	369/109
4,315,269	2/1982	Bloom et al.	346/145.1
4,378,571	3/1983	Handy	358/446
4,463,380	7/1984	Hooks, Jr.	348/580
4,568,941	2/1986	Thomas et al.	342/185
4,688,203	8/1987	Koishi et al.	369/148
4,745,475	5/1988	Bicknell	348/442
4,754,279	6/1988	Cribbs	342/185
4,833,475	5/1989	Pease et al.	342/185
4,837,579	6/1989	Pease et al.	342/197
4,845,501	7/1989	Pease et al.	342/185
4,868,653	9/1989	Golin et al.	358/133

4,899,220	2/1990	Bazile	348/584
5,006,936	4/1991	Hooks, Jr.	358/335
5,036,326	7/1991	Andrieu et al.	342/176
5,049,886	9/1991	Seitz et al.	342/26
5,051,734	9/1991	Lake, Jr.	340/700
5,111,306	5/1992	Kanno et al.	358/403
5,126,747	6/1992	Ren et al.	342/185
5,276,866	1/1994	Paolini	358/143

## OTHER PUBLICATIONS

Service Manual for LVR-5000A, Sony Corp., p. 1.  
 Service Manual for LVS-5000A, Sony Corp., p. 1.  
 Service Manual for TQ-3032F, Panasonic, p. 1.  
 Andrew Cannon, Storage and Management of Still Pictures and Electronic Graphics Published in Image Technology (Journal of the BKSTS), 72 (1990) Jun., No. 6, London, GB.  
 H. M. Morris and R. H. Orth, Image System Communications Published in IBM Systems Journal, 29(1990)No. 3, Armonk, N.Y., US.  
 K. Mun Seong and L. T. C. Fred Goeringer, An Image Management and Communications (IMAC) System for Radiology Published in Medical Progress through Technology, 18(1992)No. 3, Dordrecht, NL.



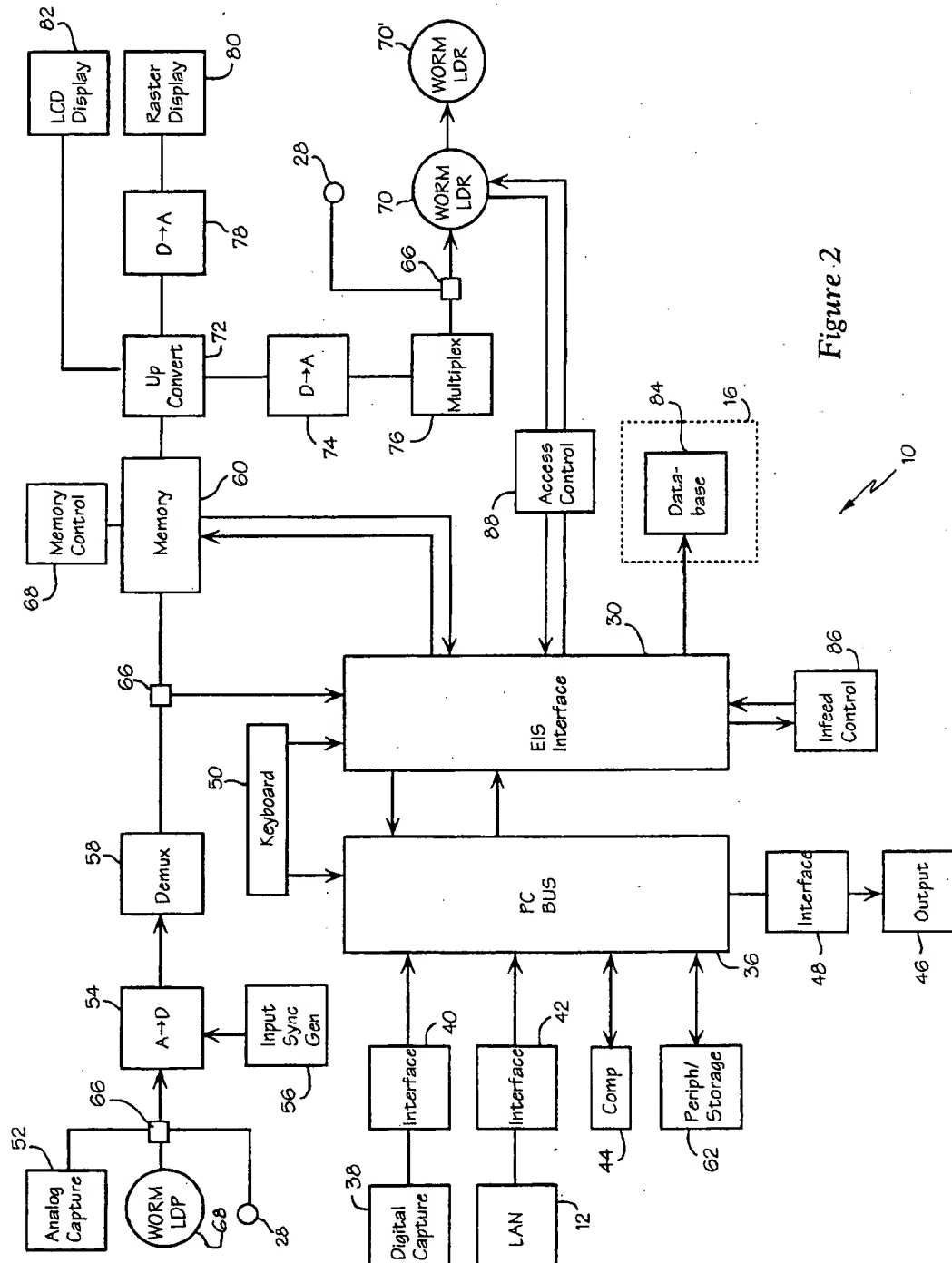
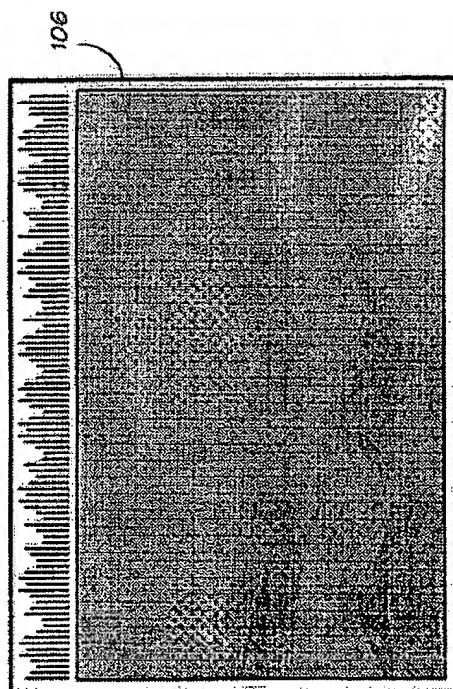
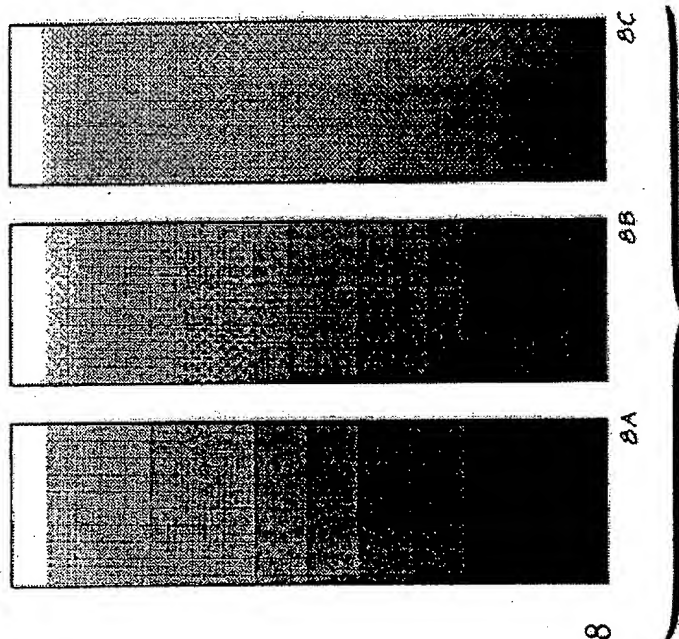
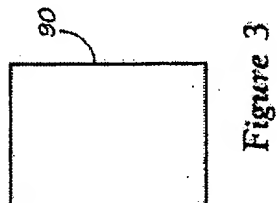
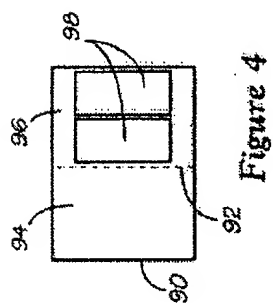
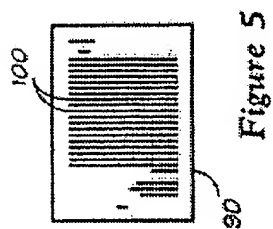
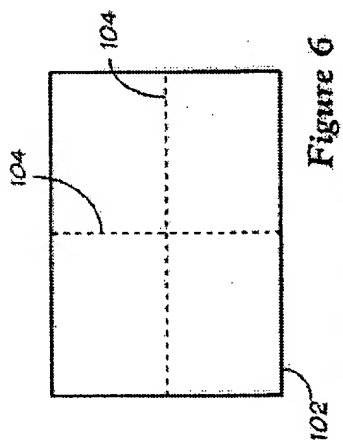


Figure 2



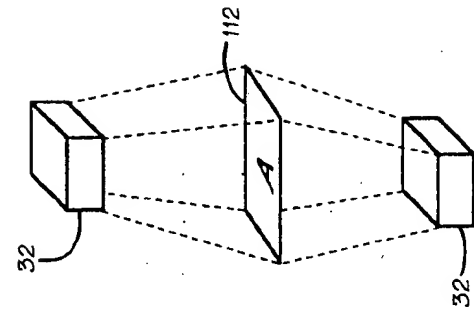


Figure 9

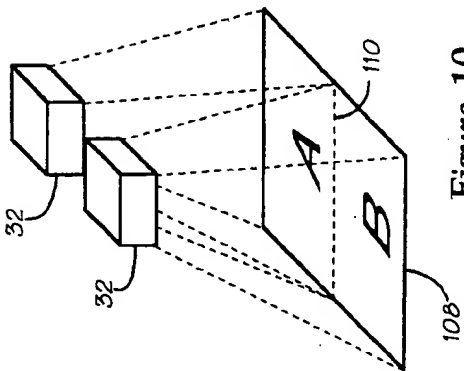


Figure 10

Figure 11

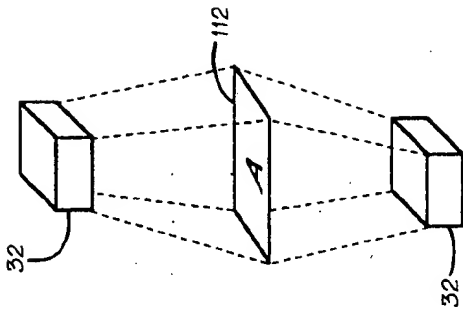


Figure 12

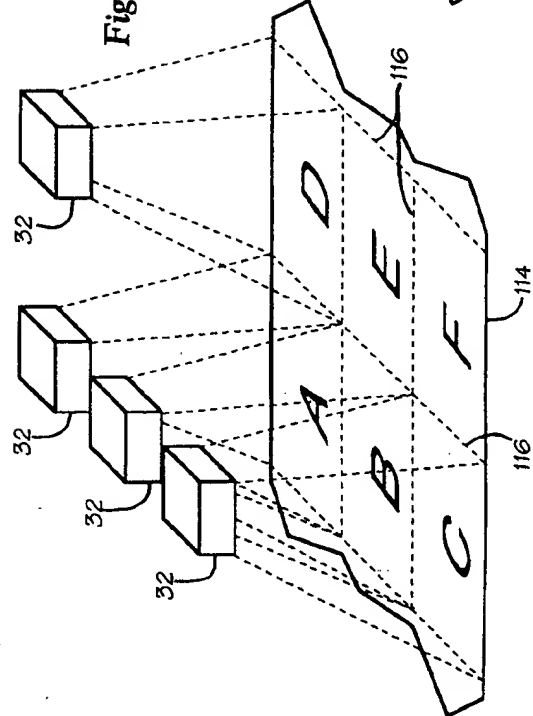
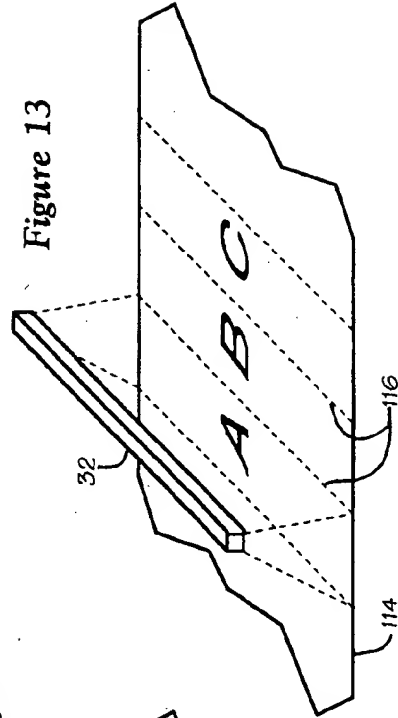


Figure 13



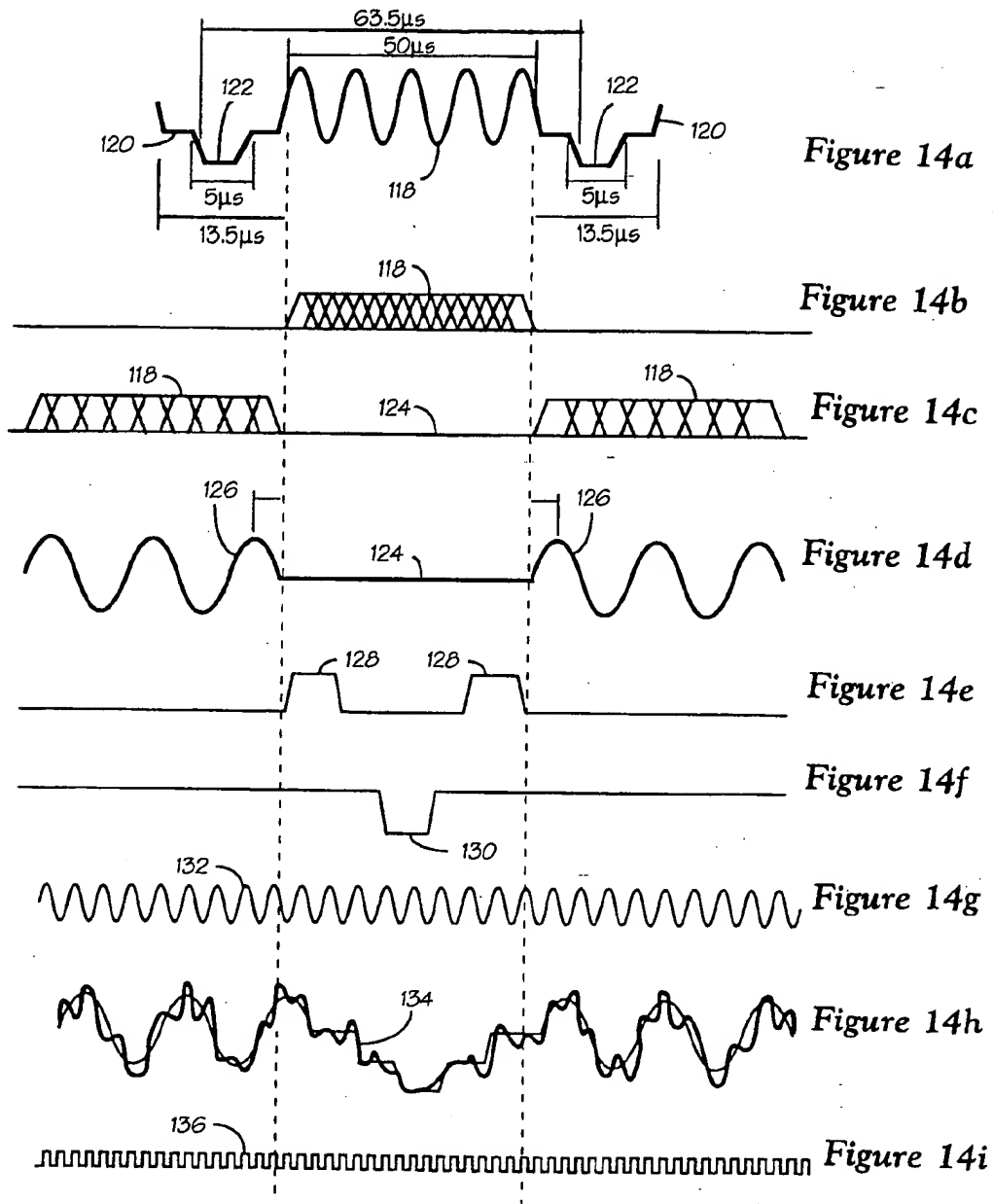


Figure 14

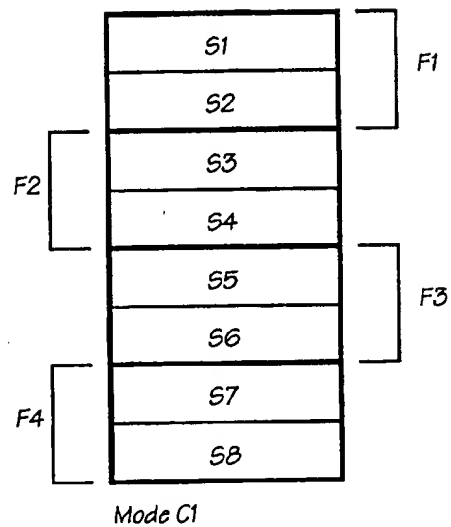


Figure 15

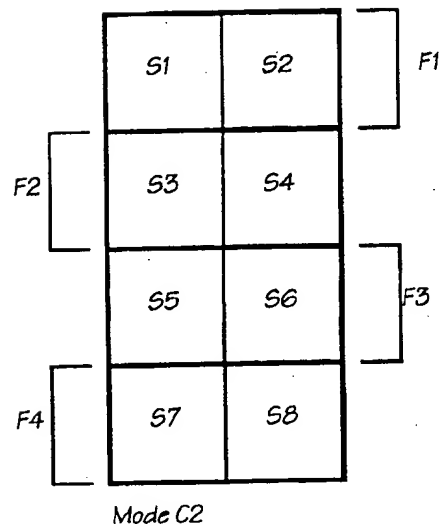


Figure 16

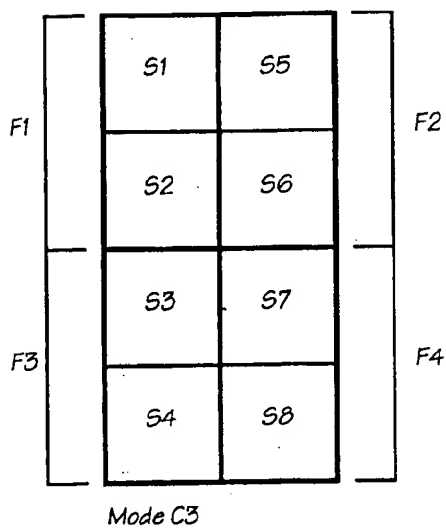


Figure 17

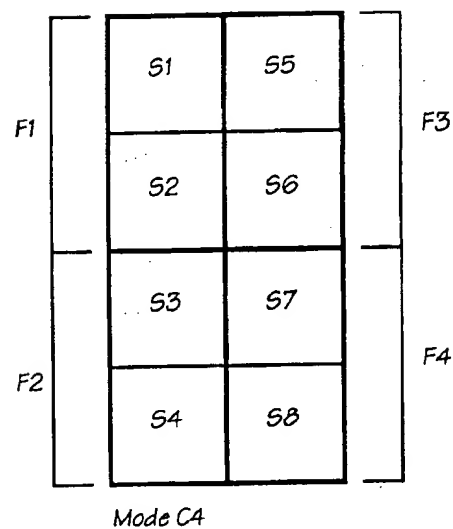


Figure 18

# METHOD AND APPARATUS FOR THE TRANSMISSION, STORAGE, AND RETRIEVAL OF DOCUMENTS IN AN ELECTRONIC DOMAIN

## BACKGROUND OF THE INVENTION

### 1. Field of the Invention

This invention relates generally to interactive document processing in an electronic domain, and particularly to a new modality incorporating hybrid (digital and analog) signal processing for the transmission, storage, and retrieval of documents to optimize informational content and processing time.

### 2. General Terminology

While the terminology utilized in the detailed description of the preferred embodiments is basic to those of ordinary skill in the art relating to designing equipment of the types discussed herein (i.e., television cameras, optical laser disc recorders, personal computer interfaces, local area networks, data transmission systems), the general terminology relating to interactive document processing operations is uncertain. A clear and uniform delineation of this terminology is therefore considered necessary to a proper understanding of the subject matter of this invention.

A document is any information recorded in a substantial (tangible) or insubstantial (electronic) form structured for human visual comprehension either directly or through the aid of enabling equipment. Documents were at one time considered written or tangible information, and this definition was later broadened to information fixed in any tangible medium such as magnetic tapes or optical disks. Currently, it is necessary to view documents even more broadly, since they may be created and reside exclusively in electronic memory, and may be transformed, presented, or transmitted without ever being reduced to a tangible medium.

Documents are therefore one subset of all the types of informational bundles, with non-documents generally being referred to simply as recordings. Documents and recordings may coexist in the same tangible medium, such as a magnetic video tape having both an audio recording coupled with a sequence of visual frames, or they may exist in related or completely separate and distinct mediums.

The differentiating requirement imposed upon documents is that they be capable of direct or assisted visual comprehension. For example, a page written in braille is normally structured for both visual and tactile comprehension. The fact that it is intended for tactile comprehension does not defeat the fact that it is also structured for visual comprehension. (In this instance, comprehension does not imply recognition or interpretation of the braille characters or their assigned meanings, but merely acknowledgement or appreciation of their existence as having content and structured form in the document which may be visually observed.) Several types of informational bundles that are structured for comprehension only by senses other than visual or auditory do exist or have been postulated, but no uniform or concise body of terminology has been developed to classify or categorize their nature or properties. Furthermore, only recordings and documents have developed as recognized means for information processing and interpersonal communication.

It may be appreciated that where documents are integrated with recordings, the term "recordings" has been adopted as the descriptive or identifying name for both components of the informational bundle predominantly because of the

particular nature of the existing technologies and mediums currently employed on a standardized basis. It remains to be seen whether this tendency will change with the proliferation of multimedia processing (that is, processing both documents and recordings in an electronic domain as integrated but separable components of an unitary informational bundle.)

Although documents existed and were used for interpersonal communication long before recordings, the technological improvements in recording over the last century have so overshadowed the improvements in processing tangible documents that the basic recording technologies were applied extensively and uniformly to both tangible and electronic documents. The current methods utilized for processing both substantial and insubstantial documents have therefore remained conceptually static since the introduction of electronic documents, due in large part to the limiting effect caused by the recent revolution in semiconductor memory and the correspondingly complete acceptance of a digital standard controlling both quantitative and qualitative precision for all insubstantial documents.

A document may exist within any one of three domains: tangible (paper, microfilm or fiche, photographic negatives or prints, etc.); electronic image (two dimensional bitmaps or arrays consisting of rows and columns of pixels each having a particular informational depth describing a monochromatic, grayscale, or color value); and electronic content (alphanumeric data strings, formatted text, spreadsheets, machine-readable programs, etc.)

There have traditionally been considered six basic operations that may be performed in association with a document: creation, reproduction, transformation, display, communication, and storage and retrieval. These operations may be performed on any document regardless of the domain in which it resides. The commonly accepted definitions for each these operations were distilled and articulated concurrently with the development of the existing technologies for converting tangible documents to electronic images, and for converting portions of electronic images to electronic content.

As may be seen by the following discussion, recent developments in the field of document processing (and particularly the technology disclosed in this application) necessitate expanding this list and refining the definitions of those operations. An appreciation of the interrelationships among and distinctions between these operations and their intended definitions is therefore a prerequisite to understanding the modality disclosed herein.

The definitions of "creation" and "reproduction" as used herein differ only slightly from the traditional protocol.

Creation is the initial authoring and fixing of a document in a specific medium. Creation can then be said to comprise the interrelated steps of composition and recordation in which the document is given content and form, both of which may be dependent on human perception and physical limitations of the medium.

Reproduction is the recording of a document's current content and form on multiple instances of the document's current medium.

At this point it is necessary to diverge more significantly from the existing protocol to interpose distinct and broader operational terms. As discussed subsequently in greater detail, current technology has produced expectations regarding operational precision in document processing that have focussed principally on the qualitative precision in reproduction of tangible documents and electronic images and the

quantitative precision in transmission, storage, and retrieval of electronic content. These expectations are no longer valid, particularly when discussing the qualitative precision associated with document processing in the electronic image domain, and to be accurate the terms must therefore reflect the fact that certain processes affect either the content or form of a document to the degree that those processes must be reclassified as different operations.

For example, photocopying has traditionally been thought of as reproduction since it is the production of a "duplicate" of an original image on a similar tangible medium. This duplicate image retains sufficient qualitative precision in both content and form that it may be utilized for interpersonal communication in place of the original for many legal or business practices. However, current black-and-white and color photocopying processes result in the loss of such a substantial amount of information in some situations that it constitutes a transformation or material alteration in the basic content of the original document compared with the levels of qualitative precision established by the technology disclosed herein.

As such, for purposes of this discussion photocopying is actually the creation of a distinct derivative document based upon the form and content of the original. In some instances, due to the nature of the original document and the photocopying technique employed, the derivative document will retain sufficient information and qualitative precision to constitute a reproduction. However, a photocopy will not in all cases be a reproduction.

Representation is the recording of a document's current content and form on a different or distinct medium, or recording a portion of the current content and form on the same or different medium, which results in a change in the informational composition of the document. Representation may therefore be thought of as involving some intermediate transition in the domain, content, or form of a document. Representation will frequently entail the transition between two domains, such as the printing of an electronic image on a tangible medium or rendering the image as electronic content, but as with reproduction the original document may continue to exist and reside in the same domain.

Because documents are more frequently being created in the electronic image and electronic content domains, it is important to remember that many people incorrectly regard the initial tangible representation of a document as the original. Those representations are in fact only derivative documents which do not possess the same informational composition as the original, and many representations of an original document result in transitions or transformations that are unintentional from the viewpoint of the operator, but which are intentional and necessary from the viewpoint of the designer of the technology being utilized to produce the representation.

Thus, any intermediate transition between domains or mediums is assumed to encompass some form of a representation, unless the particular nature and character of the original document and the processes used for the transition are sufficiently precise and compatible for the representation to be considered a reproduction for the given functional purposes being considered.

Transformation is the change in content, form, medium, or domain of a document that produces a new or derivative version which is itself a unique document. Transformation can easily be thought of as the creation of a new document, and conversely the process of creation can be thought of as a series of transformations eventually resulting in a docu-

ment having a desired informational content or form within a specified medium. In some cases (such as manipulating an electronic image contained in semiconductor memory) the original document ceases to exist and is replaced by the new or derivative document. It should also be remembered that other operations such as representation may inherently produce or require transformations due either to the technology employed or the limitations imposed by transitions between mediums or domains, and those transformations and manipulations are frequently transparent to or not appreciated by the operator. As with representation, transformation involves some intentional transition in the document invoked by the operator or the technology designer. The term manipulation is therefore regarded as being more appropriate to describe transformations that are consciously made, selected, instructed, or invoked directly by the operator to intentionally affect the content or form of the document in a predetermined manner.

Transformation has traditionally included some processes involving a change in medium or domain, however since we must assume for definitional purposes that a transition to a new medium or domain may have a substantial and often deleterious affect on the actual informational content of a document, any process involving the transition between mediums or domains are also necessarily regarded as a transformation that gives rise to either a representation of the document or the creation of a new document.

Presentation is the visual manifestation of a selected portion of the content and form of a document for human comprehension. Presentation would include processes such as displaying an electronic image as a rasterized image on a cathode-ray tube (CRT), as a bitmap image on a liquid crystal display (LCD) or light-emitting diode (LED) display, or the process of projecting a visible image onto a tangible surface using an LCD, LED, or similar device. Presentation would also include other methods of projection, such as refractively projecting a visible image from a tangible document such as film, fiche, slides, negatives, or X-rays.

Transmission is the transportation of a document through space or between remote locations. Transmission is believed to be a more accurate term than communication, since a document may be satisfactorily communicated and comprehended without requiring transmission (such as by displaying or representing the document.)

One could theoretically distinguish communication by defining cognitive boundaries for each individual involved in the processes of document handling (such as the author/creator, editor/operator, interpreter/reader) and treat communication as transporting the informational content of the document between cognitive boundaries. In comparison, transmission is the physical transportation of a document through space without regard or reference to cognitive boundaries. Transmission may then be thought of as a subset of communication, but devoid of any comprehension requirements. To the extent that communication requires comprehension of information at some level, it is strictly speaking not an operation that is performed on a document, but simply the utilization of one or more document processing operations as intermediate steps in the overall process to achieve the result of interpersonal communication.

Storage and Retrieval is the transportation of a document through time in a static state in a manner permitting the selective acquisition of that individual document from its storage medium. The provision that the document be in a static state is an addition to the traditional definition of storage and retrieval, since a document may theoretically be

transported through time by holding the document in active memory without being stored and retrieved. Similarly, the provision that the document be selectively acquired from its storage medium is an addition to the traditional definition, distinguishing storage and retrieval from the independent and unrelated operations of recording and replaying.

An electronic image normally exists in active volatile memory, and what is perceived by the operator as the document is actually a display of the document being repeatedly and instantaneously refreshed by information drawn from that memory, and manipulations of the displayed image are inserted into and held in that memory. Alternately, an electronic image may be swapped to a separate portion of memory such as a volatile RAM disk or cache, which simulates the existence of a magnetic or optical storage medium but at higher speeds. The process of exchanging informational bits in active memory, and the intervening holding of those informational bits, is not considered storage and retrieval. Conversely, one could consider the process of writing an electronic image to nonvolatile semiconductor memory such as read-only memory (ROM) to accomplish the storage and retrieval function, however semiconductor memory is rarely (if ever) utilized for the selective acquisition of electronic images.

The fact that storage and retrieval have conventionally been treated within the boundaries of one operation underscores the unique and reciprocal nature of those two processes, wherein storage implies the ability to retrieve selected documents individually and non-sequentially from among a plurality of sequentially indexed documents. The nature of document storage may be contrasted with that for audio recordings, where there is no reciprocal function corresponding to retrieval. Instead, recordings are either reproduced or they are replayed in a manner corresponding to presentation or representation. Thus, for most purposes the "storage and retrieval" function for recordings forms a closed loop through various other operations, and is normally not treated as a reciprocal function or single operation. Electronic images may be processed through very similar closed loops involving other operations than storage and retrieval when those documents are being treated as (or form portions of) recordings, such as in the case of multimedia processing applications, but these closed loops are not deemed to be storage and retrieval.

The execution of one or more of the operations described above is termed "document processing," which is considered to be a subset of "information processing" since information can exist in forms other than documents. There are similar sets of operations applied to processing information in fields other than the three document domains, each operation having a definition which may be unique or peculiar to that field. However, there is frequently some overlap between the terminology used in document processing and other fields of information processing, as well as the informal or nontechnical use of the same terminology, which can result in some inaccuracy and inconsistency.

For example, electronic content is frequently but erroneously referred to as "data" because people equate that term for computer files with the discrete elements of visual images that they recognize as conveying exact or immutable information. A person recognizes certain elements (such as numbers, letters, or symbols we term "characters") with a degree of precision that corresponds to their appreciation of how computers read data, without reference to the conceptual principles involved with communicating, understanding, or comprehending the content of that information apart from the incremental elements. A person can view a page of

text and differentiate discrete elements such as characters, each of which have an assigned meaning to that individual. The informational content of the page is considered relatively precise or exact, since it is presumed that individuals to whom the document is directed will assign the same meanings to each element. For purposes of determining precision or accuracy, we disregard both the fact that the aggregation and interpretation of those discrete elements may convey completely different concepts to each individual, and we dispense with individuals who assign different meanings to those informational elements than do members of the target group. In the same way, people understand computers to read data with exact accuracy independent from the need for recognition or interpretation, and there are basic hierarchical groupings of formats, programs, languages, and machine architectures that define what meanings will be assigned to discrete data elements for certain purposes. This appreciation may change as the general public becomes more aware of and familiar with the technological processes involved in the conversion of electronic images to electronic content through optical character recognition (OCR), the application of artificial intelligence or fuzzy logic to visual object recognition in robotics systems, or the use of electronic photography to store electronic images on optical recording medium in place of film negatives.

Electronic content can more properly be defined as encoded data plus a set of structural linkages. The data or information is encoded so that it may be utilized or operated on directly without human or artificial intelligence being applied to "interpret" or "recognize" specific information embedded within the document. The structural linkages are usually defined as accepted information storage or interchange formats and vary in complexity from simple linear data strings in which numbers are stored in a one-dimensional forward-reading sequence, through formats such as rich text format (RTF) or symbolic link format (SYLK) files representing data and particular visual attributes for displaying or representing that data on a page, to page description languages such as Postscript in which an electronic image is reduced to and expressed exclusively as mathematical formulas, callable subroutines, or vectors representing individual components from which the electronic image can be interpreted and reconstituted for display or representation.

Two distinct standards have developed for judging the threshold for integrity and requisite precision in the transmission, storage, and retrieval of electronic content versus electronic images.

The integrity of electronic content is based on a quantitative threshold, and complete or error-free precision is frequently presumed. Lack of quantitative precision may have varying degrees of impact on the qualitative integrity of electronic content. For example, a single discrete or one-bit error in an RTF file may have only a minor effect on the textual representation of a character, and therefore a negligible effect on the qualitative integrity of that file when displayed or represented as a page; the same numerical error in a SYLK file may disrupt a localized portion of the matrix such as a row or column in a spreadsheet, but leave a large portion of the spreadsheet intact and unaffected from a qualitative standpoint; a one-bit error in a program file could be fatal to the program's operation and therefore completely destroy all qualitative integrity. Consequently, preventing even one-bit errors in the transmission, storage, and retrieval of any electronic content document is assumed to be mission critical because there is an expectation of absolute quantitative precision (and therefore complete qualitative integrity) associated with electronic content.

Conversely, relatively significant quantitative "errors" may be introduced into electronic images and yet be accepted, because they leave the content and form of the electronic image effectively intact without approaching the threshold established for qualitative integrity in the electronic image domain. Again, it should be noted that thresholds for qualitative integrity in the electronic image domain are currently set at artificially low levels because they are being established or judged as a function of the human visual comprehension of presentations (raster or LCD displays) and representations (paper or film printouts) of the corresponding image.

As discussed below, the conventional wisdom in interactive document processing is to capture an image from a source document, and to keep that electronic image exclusively in the digital domain once it has been digitized. It is commonly believed that maintaining the digital nature of the electronic image preserves the accuracy and integrity of its informational content. In fact, this is a mistaken assumption.

First, it must be remembered that any digitization process relies upon an initial transition from the analog to digital domains, and there is an equal probability of introducing errors in informational content at that step as there would be for any other subsequent analog-to-digital conversion. That is, once an image is digitized it is relatively simple to verify that the "data" content does not subsequently change, however there is no assurance that the original "data" accurately reflects the true informational content of the source document. Added to this is the problem that in order to obtain reasonable transmission and storage times for most document processing, images are routinely captured and processed at bit-depths much less than 8-bits grayscale, and often as monochromatic images of very low resolution. Thus, a great deal of informational content is being intentionally discarded for no reason other than to facilitate subsequently digital processing techniques. The quantity of "errors" that might be introduced by repeated analog-to-digital and digital-to-analog conversion of an electronic image still remain many orders of magnitude below the quantity of "errors" that are interjected into images by current document processing technologies.

Second, if the capture device defines the rate limiting operation in the signal conversion process, the technology disclosed herein provides informational integrity equivalent to exclusively digital processes. If the storage device defines the rate limiting operation, and the capture/storage/retrieval/transmission pathway involves four conversions between analog and digital signals, the system has only increased the probability of losing a "pixel" by three times when compared to an exclusively digital process. Since the informational content is several orders of magnitude higher than a conventionally processed digital image, a three times increase in the probability of an error in one "pixel" is insignificant. It should also be remembered that this discussion focusses on the complete loss of a pixel, whereas in an analog frame of reference an error might actually amount only to shifting a pixel to a slightly higher or lower gray level. Isolated errors of this type would be completely imperceptible when imbedded in an image comprising more than a million adjacent pixels, with each pixel having two hundred or more possible discrete grayscale values. The actual effect of an error of this type is further reduced when one considers the potential for applying oversampling and signal averaging techniques for each analog-to-digital conversion.

The application of hybrid processing may thus be equated with a tradeoff between the sanctity of bit-value integrity for

the sake of vastly increased processing speeds. However, the loss of bit-value integrity is relatively meaningless when applying this technology, because the techniques for maintaining reasonably high precision in analog-to-digital conversions reduces the frequency at which errors may occur, and any error becomes statistically insignificant if one considers the many-order increase in the magnitude of informational content when using greater bit-depths that effectively dilute any error.

The use of 8-bit grayscale is preferable for most current interactive document processing operations since it provides a displayable image having far superior informational content than conventional techniques for storing tangible source documents (which often treat the original source document as line art and rely on conversion to monochromatic or one-bit levels which discards the majority of the document's actual informational content.) It has been shown that an 8-bit grayscale image is actually more readable than a lower level or monochromatic image corresponding to the same document, due primarily to the wealth of informational content situated in the zone situated above the full content of a lower level image and below the portion of the 8-bit grayscale image that is beyond the limits of visual perception. As such, there is a large portion of a true 8-bit grayscale image that cannot be visually perceived without the aid of enabling equipment, and which would be discarded in normal viewing. The ability to display an electronic image in 8-bit grayscale therefore also permits a wide range of user-definable adjustments in the contrast levels and grayscale filters to be applied to the viewed image that facilitate interactive discrimination or enhancement of the informational content in the image for certain applications (such as diagnostic review of X-rays, MRI scans, or other medical images) without affecting or altering the source document or its stored image.

In addition, the 8-bit grayscale electronic image will contain background details such as security paper patterns, watermarks, illegible color inks, markings that are faded or too faint to be perceived by viewing the original source document, as well as creases, smudges, stains, and other unique identifying details that assure far greater certainty when verifying the integrity and authenticity of the electronic image, far exceeding the currently accepted standards for duplicates of financial, business, and legal records. The use of 8-bit grayscale also permits the capture of electronic images from damaged or aged tangible source documents such as burnt papers or faded microfiche that could not be reproduced by other means, and which may be accomplished as if the original source documents were undamaged.

Because electronic images currently exist as bitmaps or digital arrays in semiconductor memory, they are transmitted, stored, and retrieved in digital form and these operations do not themselves introduce quantitative errors in the content of the electronic images. The most significant "errors" or losses in content generally result from intentional transformations occurring during processing steps such as digitizing a tangible document into an electronic image, or formatting an electronic image for integration into a program or as part of a file storage protocol.

It must be remembered that the acceptable threshold for qualitative integrity of electronic images as used in reference to the technology disclosed herein is several orders of magnitude greater than can be perceived by human vision when viewing a tangible document, and is at or near the limits of what can be practicably achieved by conventional representation devices such as laser printers or film record-

ers. Consequently, while the technology disclosed herein will normally be implemented in embodiments which provide less than the absolute or complete quantitative precision associated with conventional digital modalities for purely practical reasons, the qualitative integrity of documents created in or converted to the electronic image domain and subsequently transmitted, stored, and retrieved by the disclosed modality will exceed that currently accepted for document processing in the electronic image domain and yet permit substantial decreases in the time required to complete those processes.

The methods disclosed herein may be applied equally to documents in either the electronic image or electronic content domains. However, the use of this modality for processing electronic content will be only be adopted if a sufficient threshold level of quantitative integrity (as defined by the particular application involved) can be consistently maintained or verified. The complexity or expense associated with assuring this requisite level of quantitative integrity for electronic content may be commercially prohibitive given the adequacy of conventional systems now used extensively for digital transmission, storage, and retrieval of electronic content, despite the significant differential in speeds at which the technologies would operate.

Furthermore, at present the vast majority of personal and business communication is conducted using documents that remain almost exclusively in the tangible or electronic image domains, and the immediate need for applying this modality to electronic images far supersedes the comparatively insignificant demand for accelerating the processing of documents in the electronic content domain. Application of this modality to documents created in or converted to the electronic image domain provides an effective precision that is functionally indistinguishable from complete qualitative integrity, and therefore substantially greater than basic levels of qualitative integrity now utilized for tangible and electronic image documents. In addition, the complexity and expense associated with this modality are no more than for existing technologies.

For those reasons, the remainder of this discussion will focus on documents in the tangible and electronic image domains, however it is understood that the modality may be readily applied to the electronic content domain if adapted or augmented to provide suitable assurances that acceptable quantitative precision can be consistently maintained or verified throughout the transmission, storage, and retrieval processes.

Since the focus of this discussion is on particular embodiments designed for the tangible and electronic image domains, it is presumed that any "data" embodied within a tangible or electronic image exists as a function of the document's content and form. Therefore, any data contained in a document is subject to comprehension and recognition by the visual inspection of presentations or representations of the image by humans, who may then manually transcribe or encode that data for use as electronic content, or by the application of artificial intelligence to recognize, interpret, and encode that data from within the image.

There have traditionally been four additional operations associated with transposing or bridging a document from one domain to another: scanning or capture (tangible to electronic image); recognition (electronic image to electronic content); rasterization or bitmapping (electronic content to electronic image); and output or marking (electronic image to tangible.) These operations do not encompass two possible transitions which could occur directly between the

tangible and electronic content domains, however virtually all technology now in use relies on some intermediate transition through the electronic image domain.

This terminology can be somewhat misleading. For example, a raster is conventionally defined in electronics as a uniform rectangular pattern of scanning lines having an aspect ratio determined by horizontal and vertical synchronization and timing (or blanking) pulses that is produced on a cathode-ray tube (CRT) in the absence of a modulated signal. In image processing, however, a raster usually means the display of the digital array associated with an electronic image on a raster device such as a monitor, which could as easily be displayed or projected directly as a bitmap using an LCD or LED device. Furthermore, the scanning lines in a raster have no relation to the process of scanning a tangible document in most devices that are called scanners, which conventionally incorporate line- or area-array CCD technology.

It may be readily appreciated that an electronic image exists as a digital array in memory and does not require being displayed as a raster or bitmap, but such a presentation is merely an aid for human visualization and comprehension of the electronic image as it resides in memory. The display is therefore a "virtual" document and the bitmap or digital array is the true or "original" document. The process of transposing a document from the electronic content domain to the electronic image domain really constitutes mapping the image into a digital array in memory, thus the term "bitmapping" has been added to the conventional nomenclature for this transition.

Similarly, transposing a document from the tangible domain to the electronic image domain requires the same mapping of an image into a digital array in active memory, and could just as well be termed "bitmapping." In the field of document processing, where an operator works at a computer or workstation, the term "scanner" has traditionally been applied to a peripheral capture device which creates a digital array or bitmap of a tangible document, and that digital array is simply "dumped" or swapped into a segment of active memory within the computer. Conversely, if the peripheral capture device produces an analog output of sequential frames, the transition to a digital array or bitmap may be performed by the process commonly called "frame grabbing" either by the peripheral device or on board the computer. In this case, the term "capture" is utilized to include both the processes of scanning and frame grabbing where a digital array or bitmap of an electronic image is produced and resides in active memory.

As previously noted, the transition between any two domains almost always results in some transformation of the original document to a new or derivative document, whether or not that transformation is visually perceptible. Similarly, any presentation or representation of an electronic image either produces a virtual image or creates a new tangible document. Theoretically, the virtual image and the new or derivative document should be identified and treated as new documents having different informational content and form than the original document residing in memory as an electronic image.

Many factors affect the degree to which the informational content in a presentation, representation, or transformed image diverges from that of the original document. Because informational content is judged as a function of visual recognition, three factors have become basic to measuring the informational content of an image: kind, depth, and density.

Kind designates the classification of the image, and for purposes of this discussion may be monochromatic, grayscale, or color. These three kinds of images encompass most or all of the visually perceptible documents. At the same time, it should be remembered that there are other types of documents (and certain information within otherwise visible documents) that may only be perceived with some type of enabling technology. Infrared and ultraviolet represent two familiar examples where enabling technology produces images containing informational content that is not otherwise visually perceptible, but it should be remembered that true grayscale also contains large quantities of informational content that may be otherwise discarded in human visualization.

Depth is a digital measure of the quantity or "bits" of information associated with each informational bundle or picture element ("pixel.") The most frequently used depths in interactive document processing are one-bit (effectively monochrome), 4- and 8-bit for grayscale, or 8-, 24-, and 32-bit for color.

Density is the physical spacing of informational bundles in pixels per unit measurement. Density is irrelevant to an electronic image, and only becomes a factor to consider when presenting or representing an image. Density is often interchanged with resolution, and different standards and references have developed for display resolution and printing resolution. However, resolution is really a function of image comprehension as determined in the visual frame of reference. When the term resolution is used in interactive document processing, it is being used as shorthand for the "absolute resolution" which is the minimum separation between pixels or informational bundles that may be distinguished or resolved.

Density and resolution are important terms in document processing because they permit operators to specify an acceptable level for representing or presenting a document. They are also easily confused when comparing representation resolution with presentation resolution. For example, referring to FIG. 8, one might specify outputting an electronic image such as a continuous black-to-white 8-bit grayscale gradient to tangible form on a 300 dots-per-inch (dpi) laser printer for one use, but a much higher resolution for another. These specifications are further complicated by the fact that most tangible output devices are monochromatic, and grayscales are simulated or approximated by applying a selected halftone screen to the image, or printing the image as a dither pattern to approximate levels of grayscale. The halftone screens are usually denoted by the frequency (number of lines per inch) and the angle of orientation. In any case, the true resolution of the output device remains constant while the effective density of the image changes, and the actual informational content of the outputted document decreases compared to the original electronic image. A four bit grayscale gradient would therefore theoretically contain 16 shades of gray including black and white, but if printed at 300 dpi would show only about 12 gray levels due to the type of dithering pattern used and the processor's calculation of the optimal number of steps to create a smooth blend or transition between levels. As such, in the example recited above, the 8-bit grayscale gradient printed at 300 dpi resolution may result in approximately 58 visibly discernable gray levels or less (FIG. 8B) with the dots of the dithering pattern being very apparent, whereas the same gradient printed at a 3360 dpi with a 150 line horizontal screen will produce the same number of gray levels (FIG. 8C) but at a much higher resolution. As such, specifying a higher resolution may achieve a great increase

in informational content without increasing the grayscale depth, and increasing the grayscale depth may also greatly increase the informational content without requiring higher resolution. This may be compared with a true 8-bit image viewed on a monitor in which each "pixel" is displayed at one of 256 discrete gray levels, and which has a resolution on the order of 70-80 pixels per inch. It may be appreciated that increasing both resolution and grayscale depth will have a corresponding impact on informational content.

For purposes of creation and transformation in the electronic image domain, an image is usually treated as being composed of bitmaps, objects, paths, models, or renderings. Bitmaps are created and transformed by altering the characteristic value assigned to individual pixels within the bitmap. Objects and paths are transformed by altering either the fundamental definition of the object or path, or a characteristic attribute associated with that object or path. Attributes may be very simple or extremely complex, and attributes of paths may depend upon linkages and relationships to other paths and their attributes. Objects are generally self-contained. Models are the three dimensional equivalent of objects, but are composed of one or more assembled structural blocks. A rendering is essentially a complex bitmap created by applying attributes to a model, but which cannot be interactively transformed by altering the characteristic value of separate pixels.

Objects, paths, models, and many renderings are usually "device density" or "output resolution" dependent, meaning that they are treated within an interactive program as formulaic expressions each having a set depth but variable density, and when presented or represented they will adopt the highest density afforded by the capabilities of the presentation or representation device. For example, a simple arcuate path expressed in Postscript language will have a shape, size, and a specified value within the range dictated by the image's grayscale depth. When output on a 300 dpi printer, the image will have the same basic shape and size as the electronic image, but the printer will utilize its 300 dpi density to provide the best approximation or effective resolution of the grayscale value and path contours defined by the formulaic expression as possible. Output on a 1250 dpi printer will again have the same shape and size but higher density, meaning that the effective resolution of the grayscale level and path contours will more closely match the formulaic expression. Effective resolutions may be so low that losses in informational content are clearly perceptible, or so high that they exceed visible comprehension. The nature and use of the tangible document being represented or the virtual image being presented will dictate the preferred or acceptable effective resolution.

The three measures of the informational content in an electronic image are completely independent of the image's physical size. As a practical matter, limitations imposed by the processor speed, available memory, and storage medium in a document processing system will sometimes require reducing an electronic image's depth or density as its physical size increases. Kind, depth, and density may also be selectively manipulated to achieve a particular visible or perceptible result when an electronic image is transformed, represented, or presented.

Advances in both interactive and non-interactive document processing are evaluated according to five criteria: compatibility, transparency, decentralization, modularity, and operational capacity.

Compatibility refers to the capability of different technologies to utilize the informational content of a document.

Compatibility is now limited to storing a document in one or more predefined formats, with interactive document processing programs having the ability to access information only from specific formats. Usually, the higher the level of a program the more formats in which it will store and retrieve a document. If a format or conversion is unavailable, all or a portion of the informational content of the document will be inaccessible. For the most part, compatibility in transmission is limited to modem and network protocols for electronic content and facsimile and network protocols for electronic images.

Transparency has two definitions. At one level, transparency is the movement of documents between domains without loss of informational content. Operational transparency is the ability of an user to employ a technology without conscious consideration of the inherent transformations produced by that technology. In interactive document processing, effective transparency can be defined as the transition from the tangible to electronic image domain without a visually perceptible loss of qualitative integrity, and as equivalent access to electronic image documents through transmission or retrieval processes independent of the location of the original document and without regard to intermediate transformations.

Decentralization refers to the ability of users of communicating systems to perform the same document processing operations on the same documents, and to have the same capabilities that are available to operators at a central document processing or coordination facility.

Modularity refers to the linking of single- or multi-function devices for document processing. Modularity increases the functions performed by devices, or increases the available linkages between devices, to optimize paths through which documents are processed. For example, merging a scanner with a printer to accomplish the functions of scanning, photocopying, facsimile reception, and printing is a higher order of modularity than having four separate devices performing the same four functions. Another facet of modularity is scalability, which permits the addition (or subtraction) of a redundant peripheral device to a system to increase (or decrease) a particular operational capacity of the system without requiring replacement of the complete system. For example, the addition of a SCSI hard drive to a personal computer permits the system to be scaled upwardly to increase its information storage capacity.

Operational capacity for most interactive document processing is determined by measured capabilities or benchmarks such as processor clock speeds (in megahertz), millions of instructions performed per second (MIPS), bit depth of semiconductor memory, access times for semiconductor memory (in nanoseconds), transmission rates (in baud), storage capacity (in megabytes), storage density, and seek and read/write speeds (in milliseconds to microseconds).

#### Statement of the Problem

A dilemma has existed since the advent of digital processing for electronic images.

First, digital signal processing is inherently slow, and designers have been unable to obtain high resolution at fast speeds using digital technology. Second, analog signal processing is fast, but designers have been unable to produce and process high resolution images or provide the ability to interactively interface with electronic images in analog form. Furthermore, analog signal processing presents problems such as signal loss, degradation, noise, and distortion

that have been unacceptable for interactive document processing applications (particularly in transmission or storage and retrieval operations), and selective storage and retrieval capabilities have not been developed in the analog form.

In discussing the speed of transmission, storage, and retrieval of documents created in or converted to the electronic image domain, it is important to bear in mind that this modality is concerned only with electronic images that are expressed as a bitmap or digital array in semiconductor (or other) memory that will be utilized for the selective transmission, storage, and retrieval by an operator.

The constraints in the speed at which electronic image documents may be transmitted, stored, and retrieved result from the application of technologies designed primarily for electronic content to all electronic domain documents, and the inherent nature of the digital technologies used to process electronic content. For example, conventional compact discs or magnetic hard disk drives will have seek times on the order of milliseconds, and read/write times on the order of microseconds per digital word. This is approximately two to three orders of magnitude slower than the routine transmission, storage, and retrieval rates for the technology disclosed herein, which operates corresponding to the nanosecond time range that conventional static random access memory (SRAM) is capable of exchanging bit groups.

The routine transmission, storage, and retrieval rates of the technology disclosed herein will be positively affected by the utilization of memory having faster access times, as well as the capability to process larger digital words. The limiting factors become the speed at which sequential access semiconductor memories (such as CCDs) can be operated as the size of the digital arrays increase, and the signal exchange rates at which optical and magneto-optical storage media can perform.

#### Description of the Prior Art

The modality disclosed herein may be applied in either one of two forms: (1) the conversion of tangible documents to electronic images and the transmission, storage, and retrieval thereof; or (2) the creation or transformation of documents in the electronic image domain and the transmission, storage, and retrieval thereof. In the preferred embodiments for document processing in business environments, an integrated system having both capabilities operating in tandem is disclosed. There are two equally satisfactory views of this integrated system. If considering the first form, one could view the system as a four step process comprising the steps of source capture, image digitization and handling, image storage, and image retrieval. Conversely, in considering the second form one could view the system as comprising a core process of image handling (holding and manipulating a bitmap in active memory), image storage or transmission, and image retrieval, with the steps of source capture and digitization being ancillary or optional processes utilized only when converting tangible documents to the electronic image domain.

The majority of interactive document processing occurs in a user- or operator-interactive environment which may be referred to generally as interactive document processing (IDP). Almost all interactive document processing is performed using a computer as the interactive processor. The computer may be a dedicated device such as a word processor, a personal computer, workstation, or a computer or terminal linked to a file server or mainframe by a network.

In these systems, once an image is created in or converted to the electronic image domain as a digital array or bitmap, it remains in digital form throughout all subsequent operations even when the digital array is transmitted, stored, or retrieved.

Preservation of the digital format is maintained for two reasons. First, the electronic image is being handled as electronic content for purposes of storage and retrieval because current storage methods do not differentiate between electronic content and electronic images, and absolute quantitative precision is dictated for the electronic content including programs and data files. Electronic content is not separated from electronic images, and one device is utilized simultaneously for the storage of both electronic content and electronic images. The value of the depth for each pixel of an electronic image is therefore also stored on the available medium in digital form as though it were electronic content. This preservation of the digital format for transmission, storage, and retrieval of electronic images has remained almost exclusive and inviolate since the introduction of semiconductor memory. Only recently have other means of recording and replaying electronic images been introduced for use in multimedia processing, however these means have no applicability to the transmission, storage, and retrieval of electronic image documents.

Second, digital transmission systems provide better signal to noise ratios resulting in higher efficiency for a set precision. (Purely electronic systems are conventionally not measured according to precision, but rather by the efficiency achieved by in a system given a predetermined precision threshold and a set transmission configuration.) As such, digital transmission provides very high efficiency when the required precision is nearly absolute, but is extremely inefficient when compared to analog transmission for precision levels only slightly below absolute. Since absolute precision is presumed for transmission, storage, and retrieval of electronic content, digital transmission is used for both electronic content and electronic images.

The result has been to apply digital technology wherever possible to produce inter-device compatibility (for example, between telecommunications systems, facsimile transmission devices, and personal computers), and the demand for digital components used to transmit, store, and retrieve information has resulted in commercial devices which are readily available and inexpensive. In fact, the prevailing trend has been to adopt digital systems for many other information processing and recording systems such as audible telephone communication, compact disc recording, and digital audio tape (DAT). These digital technologies are not without many real merits and advantages that may readily be appreciated by both the technology designers and users of the particular consumer devices, and there has been a corresponding shift among product designers and consumers towards preferring digital technology for purely aesthetic reasons.

Examples of the proliferation of digital technologies for storing and retrieving electronic images include many of the CD-ROM drives and applications that are commercially available for personal computers, the development of the Kodak Photo-CD standard currently being implemented for transferring 35 mm photographs directly to compact disc, and the use of rewriteable optical and magneto-optical disc recorders for file storage.

Even though it is known that the qualitative integrity of an electronic image may not be perceptibly diminished by the utilization of an analog format, no technology has yet been

developed for the transmission, storage, and retrieval of electronic images in the field of interactive document processing.

Outside the field of interactive document processing, there are some representative examples of new developments concerning the application of analog technology to information processing.

Efforts to produce major developments in analog technologies for television and broadcasting have been focussed on the development of high-definition television systems, the production of larger array CCD cameras, the utilization of satellite transmission, and the use of optical laser discs for recording.

In the field of radar scan conversion, various processes have been applied to increase the rate at which polar coordinate maps can be processed and converted to an X-Y memory map and displayed using raster devices. U.S. Pat. No. 4,754,279 to Cribbs discloses a scan converter which accomplishes the conversion of an analog radar signal to a digital array through two high speed memory buffers and an intervening segmentation of the electronic image in active memory, and the subsequent assembling and display of that array as a raster image corresponding to the bitmapped electronic image. Similar polar to X-Y scan converters are also found in ultrasonic and other medical imaging devices, with a representative example being U.S. Pat. No. 4,275,415 to Engle.

Scan converters related to those disclosed in the Cribbs '279 and Engle '415 patents have also been utilized with certain modifications in systems to scan tangible documents and store the corresponding electronic images as frames in a conventional video tape format. These systems may be interfaced with a conventional personal computer, and representative examples of such systems include those currently marketed by Folsom Research of Folsom, California and RGB/Spectrum of Alameda, Calif.

Although the technology disclosed in the Cribbs '279 and Engle '415 patents and the Folsom or Spectrum devices permit operators to record sequences of electronic images at rates exceeding those for current digital components, they do not provide the required capability for transmitting, storing, or retrieving electronic images.

In particular, the technology related to radar scan converters is directed to monochromatic images, and involves certain operations for correlating discrete segments of adjoining raster lines to the corresponding pixels in a digital bitmap. The current technology also relies on using a nibble-mode for reading and writing information from high speed buffers to update only the pixels in the digital array corresponding to radar pixels that are activated or faded as a target moves, and this technology does not translate to grayscale or color images having far greater depth, nor to the transmission, storage, or retrieval of complete electronic images.

The commercial devices that may be interfaced with document processing systems further do not permit indexing and acquisition of selected images at workable rates for interactive document processing operations, but rather rely on replaying instead of retrieving documents. These devices utilize the broad bandwidth obtainable with magnetic tape mediums for recording electronic images having greater than one-bit depth in a format that corresponds to a video "frame." In comparison, the storage and retrieval of electronic images at high rates (in tens of images per second) and large volumes (tens of thousands of images per unit of medium) requires indexing and selective acquisition of

stored images on a medium such as an laser or optical disc, which has fixed bandwidth and limited frame composition.

Retrieval and display of stored electronic images at roughly comparable rates (one or more images per second) and volumes (fifty thousand images per unit of medium) are currently accomplished on rewriteable optical disc recorders such as the Panasonic LQ-4000 Erasable Videodisc Recorder, however this returns the storage and retrieval operation to the digital format, and current systems only permit limited operation for capture and storage before encountering downtime for processing.

The application of hybrid signal processing to accomplish signal compression for serial recording is also known to the art.

U.S. Pat. No. 5,006,936 to Hooks discloses a method of compacting a analog signal corresponding to unframed continuous-signal data output wherein an instrument's analog output signal is converted to a digital signal, and portions of the digital signal are alternately written to two memory buffers. As the digital signal is being written to one memory buffer, the other memory buffer is simultaneously read at a faster rate than it was written. The compacted digital signal is then converted to an analog signal having a greater bandwidth than the original data output signal, and the appropriate synchronization signals are applied corresponding to a conventional NTSC video format. The compressed and NTSC-synced signal is then stored using conventional video equipment such as a magnetic tape or optical laser disc recorder. A decoder circuit is utilized to strip the synchronization signals and decompress the signal back to the original unframed continuous-signal data output for serial replay at the same rate as the original data was produced. The Hooks '936 device therefore compresses a one-dimensional time-variable signal by increasing its bandwidth sufficiently that synchronization pulses may be added to fill the gap between the compressed signal and a standard NTSC video format, and then utilizing that NTSC equipment for recording and playback.

Methods for compressing analog signals using hybrid signal processing similar to that disclosed in Hooks '936 have been utilized for many years. While the Hooks '936 device provides an effective system for the recording and playback of large spans of unframed continuous-signal transmissions such as flight recorder data using lower cost NTSC video equipment, compaction methods such as disclosed in Hooks '936 are not applicable to the selective and non-sequential transmission or storage and retrieval of two-dimensional electronic images. First, the compression method described in Hooks '936 is most suitable for magnetic tape recording which may accommodate and utilize the comparatively unlimited bandwidth of video tape, whereas the bandwidths associated with the informational content of two-dimensional electronic images (particularly those having conventional grayscale or color depths used in interactive document processing) are already greater than NTSC equipment would permit, and signal compression would only increase the bandwidth. Second, compaction methods such as disclosed in Hooks '936 permit only serial playback of continuous-signal data output. Third, continuously alternating the flow of signal to or from two separate memory buffers prevents utilizing those memory buffers at the junction with an interactive document processing interface.

Various references in the prior art have discussed "stitching" techniques used to combine the outputs from two CCD arrays for use in a flatbed scanner, and for permitting the use of NTSC display devices for viewing high definition television (HDTV) broadcasts.

U.S. Pat. No. 4,378,571 to Handy discloses a technique for increasing the field of view of a flatbed scanner while maintaining the device's original resolution by using two CCD chips fixed in an array having a specified displacement relative to one another that corresponds to one half of the total field of view. The analog output signals of each CCD are sampled and averaged, and the gross grayscale levels for the two halves are equalized with one another by amplifying the DC gain level for one or both of the signals toward a common level. The signals are then processed by an analog-to-digital converter to digitize the image.

This technique is essentially a simplified version of "ganging" smaller CCD chips to produce a larger X-Y pixel matrix, wherein a first row of pixels are read sequentially from one CCD chip and the aligned first row is then read from the second CCD chip, followed by the second pair of rows from each CCD chip, and so forth to produce a "horizontal" bitmap. It is also possible to simply read all the rows of one chip followed by all the rows of the second, to produce a "vertical" bitmap. Ganging more than two CCD chips in each direction results in a combination of both horizontal and vertical bitmapping between different CCD chips. This technique was used in video cameras when CCD chips having only a 512x512 X-Y pixel matrix were available, however cameras using CCD chip matrices of 1024x1024 and 2048x2048 pixels are now readily available (and could similarly be ganged to produce even larger matrices.) The principle of the Handy '571 patent has apparently also been applied to hand-held scanners which are used to sequentially scan parallel strips of a source document, wherein the speed at which the timing wheel within the scanner is rolled across the document provides a crude physical control or reference for the number of pixel rows-per-inch that are scanned (or conversely, the ratio by which one strip is scaled along its length to match the pixel rows-per-inch of another strip.)

As such, the Handy '571 patent does not teach an actual "seaming" process for reconstituting divided images, and it also points out the inherent errors in the informational content of an image that may be surreptitiously introduced on a very large scale when using conventional digital techniques for document processing operations. The Handy '571 patent further does not teach any of the time-based correction and masking principles necessary to make the actual seaming of a divided image possible, nor to overcome time-based errors introduced by optical laser disc recorders and other devices. The Handy '571 device also places the analog multiplier used for gain correction upstream of the analog-to-digital converter used for digitization, and therefore reinforces the existing assumption that conversion to the digital domain is a priori necessary for image processing.

U.S. Pat. No. 4,899,220 to Basile discusses a technique for permitting the center portion of a wide aspect ratio high definition television (HDTV) broadcast to be viewed on a conventional NTSC display, and alternately a display apparatus used in combination with a conventional NTSC display for viewing the entire wide aspect ratio image. The Basile '220 patent teaches dividing the scan lines into segments corresponding to a center portion of a wide aspect display and two side panels, converting to the NTSC format for broadcasting, and reassembling those segments after broadcasting. However, this procedure produces visible lines vertically bisecting the display due to dot crawl. Consequently, Basile '220 also teaches randomly varying the location of the stitch points between the line segments of adjacent lines or between fields, thus scattering the stitch points throughout the image and varying their location

between fields to a degree where the stitch points will not produce a visible pattern interrupting the image when displayed at 30 frames per second.

It may be readily appreciated that Basile '220 does not teach eliminating the visible stitch points, but merely relocating them so that a viewer will not perceive them when viewing a many changing images displayed each second. As such, the technique of Basile '220 cannot be applied to document processing operations where the introduction of stitch points would still degrade the informational content and undermine the integrity of a stored image. Even though the degradation might not be visible when displaying many changing images, some degradation would be visible when viewing a static display of the same image, which would render the system unsuitable for many operations such as medical imaging. Furthermore, since the Basile '220 process relies on random distribution of stitch points among unrelated images, the effects of these degradations would accumulate rapidly when repeatedly storing and retrieving the same image, to the point where the image could be effectively destroyed over time. While the technology of Basile '220 is suitable for raster display purposes where the camera or broadcast generator has a very stable output, analog storage devices such as laser disc recorders or players have significant time-based errors in their output signals which cannot be corrected for using the technology disclosed in Basile '220, just as it does not provide any time-based correction to prevent the visible stitch points. The Basile '220 and Handy '397 patents are therefore similar in the fact that they do not give any consideration to absolute timing relative to an external reference.

Moreover, while the masking technique accomplished by varying the time at which line segments are switched in controlled amounts is a strong feature when considering a displayed broadcast image, the method disclosed in the Basile '220 patent apparently only permits manual control over image content without providing any grayscale calibration, DC matching, gain adjustment, or color control between adjacent line segments or on a line-by-line basis. In addition, the Basile '220 device (with or without the ability to display the side panels) is still limited to the conventional resolution of the display, which in this case is the normal NTSC resolution.

As such, technologies of the type disclosed in Basile '220, Handy '571, Hooks '936, and many other references in these distinct fields of art provide very good solutions to particular problems associated with converting between display or recording formats, ganging capture devices to increase the effective field of view, and the like.

However, it is believed that the prior art does not teach, suggest, or even address the much broader concept of interactive document processing in an hybrid (analog/digital) domain utilizing techniques such as multiplying the effective resolution of a capture device, segmenting the existing bitmap of an electronic image to optimize its informational content versus its transmission, storage, or retrieval times for a given operation or existing apparatus, storing an electronic image having a predetermined informational content on randomly accessible storage mediums having fixed bandwidths which are smaller than those associated with the original image, transmitting an electronic image via a transmission pathway similarly having a limited bandwidth smaller than that associated with the original image, integrating these processes in an interface which permits interactive utilization of an electronic image and transparently passing the image back-and-forth between conventional low speed digital processing pathways and

high speed analog processing pathways, nor the many independent advantages and synergistic effects achieved by applying these concepts as part of a complete document processing environment, network structure, or communications system.

#### BRIEF SUMMARY OF THE INVENTION

It is therefore one object of this invention to provide a method for capturing, transmitting, storing, and retrieving documents created in or converted to the electronic image domain which overcomes the obstacles imposed by the inherent speed limitations of the digital domain and the low resolution and perceived imprecision of the analog domain, yet which permits subsequent use of digital interfaces for interactive document processing.

It is a related object of this invention to design the above method such that it utilizes hybrid signal processing to accomplish the dissection of an electronic image into a progressively increasing number of segments and the a posteriori<sup>1</sup> reconstruction of a complex electronic image or a multi-image document from those segments.

<sup>1</sup> Literally "after the fact" or "deduced from experience;" implying going from effect to cause, or from fundamental components to complex structure, by reversing a previous process or chain of logical operations.

It is thus another object of this invention to design the above method such that it permits the unrestricted deconstruction or disassembly of a source document in order to optimize the informational content and processing time needed to accomplish a particular document processing operation using specified equipment to achieve a desired end result, and further permits the subsequent seaming together of a virtually unlimited number of related segments (or unrelated image memory maps) to reconstitute or create documents having extensive physical dimensions.

It is accordingly an additional object of this invention to design the above method such that it employs both digital and analog signal processing techniques at predetermined points to achieve the particular advantages associated with the respective signal processing domain, and minimize or eliminate the adverse characteristics of the converse domain.

It is an object of this invention to provide a system for processing documents created in or converted to the electronic image domain which treats the electronic image in a "fundamental" form independent of any transformation or manipulation, and which is therefore universally compatible with any existing or later established standard or format for digital input, output, transformation (including compression), presentation, or representation.

It is an associated object of this invention to facilitate the transparent integration of such a system into current interactive document processing environments.

It is yet another object of this invention to design the above system so as to introduce virtually unlimited scalability for capture, storage and retrieval, and transmission capacities to relatively large document processing systems and databases, particularly for networks in which the files server is a mainframe computer.

It is a unique object of this invention to design a system for the capture of electronic images from tangible source documents, and the subsequent transmission, storage, and retrieval of those electronic images, which permits complete decentralization of document processing operations among remote networks, and permits remote operators to transmit, store, retrieve, and receive documents with the same capabilities and access available to a central document processing or coordination facility.

It is a related object of this invention to increase the modularity of a document processing environment by linking storage and retrieval functions performed by peripheral devices with the processors responsible for document transformation and presentation, thus enhancing the operational capabilities of the entire system compared with existing technologies, and to further increase the modularity of the system by linking components by both conventional and high speed networks.

It is a further object of this invention to design the above system such that it may be constructed using currently available recording technologies and established frame formats, as well as subsequently developed technologies and formats for recording electronic images including but not limited to those for optical laser disc recorders utilizing blue-green lasers, greater bit depths, and multiple substrates.

It is a related object of this invention to design embodiments of the above system incorporating time-based corrections to allow the invisible seaming of segments in an image and correcting time-based errors that are introduced by conventional recording devices such as analog optical laser disc recorders, as well as incorporating calibration controls that permit "on the fly" adjustment of grayscale levels between lines or segments, maintain grayscale consistency and precision throughout all frames, and permit selective viewing enhancement by displaying a range of grayscale levels corresponding to a desired amplitude examining area within the image at the optimal base level for an existing display device and for a particular application (similar to the effects created by adjusting "brightness" and "contrast" to optimize a grayscale range within a displayed image.)

It is a specific object of this invention to design several embodiments of the above system for capturing, storing, and retrieving electronic images from within an interactive document processing network which permits electronic images corresponding to US letter-sized pages (or smaller) to be captured and stored in at least 8-bit grayscale at batch processing rates of 5 to 30 source documents per second or greater (depending upon the image size and informational content), and transmitted at rates of 18,000 to 108,000 documents per hour or greater.

It is a related object of this invention to design the above system such that it may be utilized to establish separate high speed network and facsimile transmission protocols for transmitting documents at equivalent speeds, as well as integrated with existing networks and communications systems in a manner that would increase the transmission rates which may be achieved using conventional transmission pathways.

It is an object of this invention to design the above system such that transmission operations will be compatible with traditionally non-document based communications systems, including such alternatives as multichannel RF television transmission, broadcast cable television transmissions, microwave and satellite transmission, and be compatible with conventional communications-based security techniques such as line scramblers and encryption algorithms.

It is a distinct object of this invention to design the above system such that it may be immediately implemented in existing document processing environments (such as financial institutions, medical centers, universities, businesses, and research facilities) using conventional interactive document processing hardware (personal computers, networks, and mainframes), and commercially available equipment for image capture and mass storage which adhere to existing uniform standards and formats.

It is therefore a related object of this invention to design

embodiments of the system such that they may be interfaced with conventional personal computers and local area network systems operating with existing commercial capabilities, one representative example being a personal computer having an ISA bus, an Intel 486 processor, a clock speed on the order of 33-66 Mhz, 70 nanosecond RAM, and interfaced with a conventional Novell or Ethernet network.

It is another object of this invention to design embodiments of the system wherein each image may be stored on a unit of storage medium and addressed by a volume and frame number, and wherein a field may be added to established databases in order to correlate batch-processed documents with their original source documents or transactions, or conversely a separate database may be utilized as a "look-up table" to identify or retrieve the volume and frame number of the stored image corresponding to a specific source document or transaction.

It has been proposed that breakthroughs in the operational capabilities of five technological areas will be necessary to convert the majority of existing business and technical document processing environments from "primarily tangible" to "primarily electronic," thus achieving an effectively "paperless" document processing environment (irrespective of source document form.) Those basic technologies are storage/retrieval servers comprising file storage and indexing capabilities, transformation servers to permit complete or universal compatibility of document formats, communication servers that provide transmission interfaces for both networks and remote transmissions (i.e., satellite, fiber optic, telephone), recognition servers capable of capturing and preserving the fundamental information from tangible documents, and document managers to track the location, flow, and transformation of documents.

It is therefore an object of this invention to design the method and system of this invention such that it achieves the goal of providing sufficient hardware capabilities in each of the five basic technological areas that would permit a primarily electronic or "paperless" document processing environment for business and technical applications.

It is a related object of this invention that the system facilitate the retention of electronic images corresponding to the "original" tangible source documents which contain and preserve significantly more usable information than can be appreciated or recognized by human visual perception of the source document, and which further may provide an audit trail and security capabilities to ensure and verify the authenticity of the fundamental electronic image.

It is a further object of this invention to design the hardware components necessary to establish an electronic publishing and archive system (such a system having previously been postulated to maintain every "original" document in storage on only one of a plurality of file servers that may each be accessed independently by any operator within the system, and in which distinct documents may be connected to one another using hypertext pointers) within existing time and expense parameters for users of conventional business and technical informational databases.

Briefly described, the invention comprises a method and apparatus for the high speed conversion of tangible source documents to the electronic image domain, and the subsequent transmission or storage and retrieval of those electronic images, utilizing hybrid (analog and digital) signal processing. The system employs a higher bandwidth analog signal for image capture and lower bandwidth analog signal for transmission or storage and retrieval, with an intervening digital memory buffer utilized to construct a bitmap of the image to facilitate various dissection and seaming schemes

which optimize image content and processing time depending upon the size of the source document, the informational content or "resolution" necessary for the particular application, and the particular storage medium or transmission pathway being considered.

The system is designed around a conventional bus structure, and the memory buffer may also serve as a transparent and universal junction with conventional interactive document processing systems including personal computers, networks, transmission systems, and various types of peripheral input/output devices. The system also processes electronic images in a manner that provides complete compatibility with formats and standards for digital storage, image transformation and compression, local area networks, and communications systems.

In a representative embodiment, the electronic image corresponding to a tangible source document such as an 8½"×11" page is captured using a camera producing an analog output signal with conventional raster synchronization. The vertical and horizontal synchronization pulses are stripped from the analog signal, which is then digitized in 8-bit grayscale and multiplexed to the memory buffer where the image exists as a digital array or bitmap that may be divided into a plurality of segments. The digital content is read from the memory buffer according to this dissection scheme, converted to an analog signal, and control signals are added. The control signals include horizontal and vertical synchronization pulses (and interval blanking), a pilot signal to maintain alignment along the seams between adjacent segments of the electronic image when it is reconstituted, and calibration pulses to permit instantaneous adjustment of the base grayscale level for each line of the electronic image to maintain a consistent grayscale level. The resultant analog signal is stored on a randomly accessible storage medium such as a conventional analog optical laser disk recorder (LDR) as one or more frames (each frame corresponding to a segment of the electronic image), or the resultant signal may alternately be transmitted to a remote location and reassembled, or up converted and displayed on a conventional monitor.

The primary advantages of the technology disclosed herein are therefore speed, increased qualitative integrity, enhanced image content, and universality or translatability of image content. The most compelling rationale for application of this modality in present business or commercial environments is the increased speed in document processing. There is a lesser demand to increase or enhance image content or qualitative integrity, however the inherent increase in stored image content that may be achieved without increasing storage densities or introducing undue expense or complexity to the overall system does have the effect of making this modality optimal for processing certain tangible documents as electronic images (such as written records of financial transactions, medical imaging, and damaged or unreadable originals) where preservation and reproducibility of all fundamental information in the tangible document and the corresponding certification of security and authenticity permit electronic document processing to be applied where it was previously too costly, complex, volatile, or susceptible to challenges of unfitness or impropriety.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a diagrammatic depiction of an interactive document processing network utilizing the electronic document transmission, storage, and retrieval system of this invention for the capture, transmission, storage, and retrieval

of electronic image documents;

FIG. 2 is a schematic diagram showing the components of an embodiment of the electronic document transmission, storage, and retrieval system of this invention particularly adapted for the capture, transmission, storage, and retrieval of electronic image documents;

FIG. 3 is a diagrammatic depiction of an 8½"×11" tangible source document oriented in the upright or vertical position;

FIG. 4 is a diagrammatic depiction of the source document of FIG. 3 oriented in the sideways or horizontal position, with a phantom line showing the source document divided corresponding to two horizontally adjacent image segments, and further showing two sub-sections corresponding to the front and back of a conventional negotiable instrument (check) superimposed over the right segment of the source document;

FIG. 5 is a diagrammatic depiction of the source document of FIG. 3 oriented in the sideways or horizontal position showing the orientation of typed lines;

FIG. 6 is a diagrammatic depiction of a source document such as a C-size drawing, with phantom lines showing the source document divided into four horizontally and vertically adjacent regions corresponding approximately in size to the source document of FIG. 3;

FIG. 7 is a diagrammatic depiction of a source document (such as a seismic tracing) corresponding approximately in size to several of the source documents of FIG. 3;

FIG. 8 is a diagram showing three grayscale gradients in which 8A is a 4-bit grayscale gradient printed at 300 dpi showing 12 gray levels including black and white, 8B is 8-bit grayscale printed at 300 dpi resolution showing the dithering pattern appearing as dots of progressively increasing diameter from top to bottom, and 8C is the same 8-bit grayscale printed at 3360 dpi with a 150 line screen showing the same dithering pattern with approximately 58 gray levels appearing as horizontal banding;

FIG. 9 is a diagrammatic depiction of a single capture device utilized to capture an image associated with the source document of FIGS. 4 or 5, with a phantom line showing the source document divided corresponding to two horizontally adjacent image segments;

FIG. 10 is a diagrammatic depiction of a plurality of capture devices utilized to capture an image associated with a source document greater in size than the source document of FIG. 3, with a phantom line showing the source document divided corresponding to two adjacent image segments;

FIG. 11 is a diagrammatic depiction of a plurality of capture devices utilized to capture an image associated with both the front and back faces of a source document such as the negotiable instrument (check) shown in FIG. 4;

FIG. 12 is a diagrammatic depiction of a plurality of linear array type capture devices configured in an array utilized to capture an image associated with a source document greater in size than the source document of FIG. 3, with phantom lines showing the source document divided corresponding to an equal plurality of adjacent image segments;

FIG. 13 is a diagrammatic depiction of a line array type capture device utilized to capture an image associated with a source document, with phantom lines showing the source document divided corresponding to an image segment;

FIGS. 14(a-f) are waveform diagrams in which 14a shows synchronization pulses and blanking intervals on opposing sides of a segment of analog signal; FIG. 14b shows the analog signal of FIG. 14a after digitization; FIG.

14c shows the digital signal split and expanded to form two digital segments; 14d shows the two digital segments of FIG. 14c converted to analog signals; 14e shows two positive-going calibration pulses; 14f shows a negative-going synchronization pulse; 14g shows a low level pilot signal; 14h shows a composite waveform of the two active signals of FIG. 14d summed with the signals of FIGS. 14e-14g overlaid over the sum of the signals of FIGS. 14e-14g; and 14i is a square wave corresponding to the clock speed utilized to reconstruct the electronic image represented by the analog signal of FIG. 14a;

FIG. 15 is a diagram showing the arrangement of segments and frames in Mode C1 (line delay) overlying a raster image composed of horizontal scanning lines;

FIG. 16 is a diagram showing the arrangement of segments and frames in Mode C2 (field delay) overlying a raster image composed of horizontal scanning lines;

FIG. 17 is a diagram showing the arrangement of segments and frames in Mode C3 (one frame delay) overlying a raster image composed of horizontal scanning lines; and

FIG. 18 is a diagram showing the arrangement of segments and frames in Mode C4 (two frame delay) overlying a raster image composed of horizontal scanning lines.

#### DESCRIPTION OF THE PREFERRED EMBODIMENTS

The electronic document transmission, storage, and retrieval system of this invention is represented in FIGS. 1-14 and referenced generally therein by the numeral 10.

Referring particularly to FIG. 1, an interactive document processing environment is shown which includes a low speed local area network (LAN) 12 such as a Novell or Ethernet network or a peer-to-peer network system linking a plurality of nodes which may be occupied by personal computers 14, terminals, or workstations. The network 12 is optionally linked to a mainframe 16 or file server along a conventional digital communications pathway. Each computer 14 on the network 12 is linked to various peripheral devices including a digital document input device 18 such as a document scanner, an output or representation device 20 such as a laser printer or film recorder, and a transmission interface 22 such as a modem or facsimile machine. Each computer 14 or terminal preferably has various input and output devices for the operator including a keyboard, mouse, trackball, or pressure sensitive tablet, storage mediums such as a magnetic or optical disk, tape cartridge, or compact disc drive, and at least one presentation system such as a raster display or LCD projection screen associated therewith.

Each of the computers 14 in the network 12 are simultaneously connected to a separate high speed network 28 which includes an interface module (not shown) installed in or linked to each of the computers 14. The interface module will preferably have the capability for at least retrieval and transmission of documents utilizing the high speed network 28, and may optionally include storage capability. The network 12 may also be connected to a remote network 24 through a conventional transmission pathway 26 such as telephone lines or a dedicated cable system, and a "high-speed" transmission pathway 26' capable of transmitting electronic documents at rates equivalent to the faster operating speeds of the high speed network 28 of the system 10.

The high speed network 28 is linked to an electronic image server 30, which is in turn connected to a capture device 32 and a mass storage device 34. The mass storage device 34 is connected to the electronic image server 30 by

both a conventional digital communication pathway and a high speed pathway. The mainframe 16 is also connected to the electronic image server 30 along a digital communications pathway.

The electronic image server 30 may be self-contained within its own housing and include the necessary ancillary components such as a power supply, operator interface, and the like, or may be fabricated on a separate interface module mounted in a host processor such as a personal computer or workstation. It is understood that each computer 14 in the network 12 that is linked to the high speed network 28 will also include a separate EIS interface 30 which has some or all of the processing capabilities of the electronic image server 30. Thus, for purposes of clarity, the term electronic image server 30 will be used to refer to the separate unit as shown in FIG. 1 that is used primarily for capture and batch processing operations, and the term EIS interface 30 will be used to refer to a computer-based interface associated with each computer 14 or workstation.

Each EIS interface 30 may include some or all of the components necessary for storage, retrieval, transmission, or presentation depending on the number and extent of document processing operations that will be performed at that station or node. As such, the EIS interface 30 is hereafter assumed to have the full functional capabilities of, and be functionally indistinguishable from, the separate electronic image server 30 unless otherwise noted, and the schematic shown in FIG. 2 will therefore delineate the relationships of both the separate electronic image server 30 and the computer-based EIS interface 30 as being mounted on a host processor such as a personal computer or workstation.

In normal operations, images from tangible source documents will be captured using the capture device 32 and stored on the mass storage device 34. A volume and frame number for each image will be written to a database in the mainframe 16 using the conventional digital communications pathway. An operator at a computer 14 on the network 12 will issue a retrieval command to the mainframe 16 for a particular document image, and the mainframe 16 will access the database to determine the correct volume and frame numbers for the image. The mainframe 16 will issue an instruction through the digital communication pathway (and electronic image server 30) to the mass storage device 34, which will retrieve the appropriate number of frames from the storage medium in the mass storage device 34 and transmit those frame over the high speed network 28 to the requesting computer 14. The EIS interface 30 within the receiving computer 14 will reconstruct the electronic image for the operator to view, and the operator may process that electronic image as desired. Alternately, an image may be retrieved from the mass storage device 34, reconstituted in the electronic image server 30, and transmitted along the digital network 12 through the mainframe 16 and to the requesting computer 14.

If the operator working at one of the computers 14 or workstations makes a transformation to the electronic image and wishes to store that derivative image or a newly created electronic image, there are two alternatives. First, the operator could issue a store command which would cause the electronic image to be transmitted over the high speed network 28 to the mass storage device 34, with a corresponding instruction being sent to the mainframe 16 over the conventional communications pathway which would update the database with any information concerning the processing history of the electronic image, and also initiate the appropriate subroutine in the mainframe 16 to instruct the mass storage device 34 to receive and store the correct number of

frames, and obtain the volume and frame numbers from the mass storage device 34 to be placed in the database. Alternatively, the operator could send the electronic image in digital form through the mainframe 16 to the electronic image server 30 using the conventional digital communication pathways, where the image would be converted to frames and stored on the mass storage device 34 with the database in the mainframe 16 being similarly updated.

The remote network 24 will ordinarily be associated with a separate mainframe 16' or fileserver, and a separate high speed network 28' with a separate mass storage device 34' as well as at least one separate electronic image server (not shown) and the associated peripherals and components. If an operator wishes to transmit an electronic image to the remote network 24, the electronic image 24 may be transmitted over the high speed transmission pathway 26' to a high speed network 28' linking the nodes of the remote network 24, where it may be saved on a separate mass storage device 34' and a database in the remote mainframe 16' will perform the necessary cataloging and indexing functions. As discussed in greater detail below, it may be appreciated that images may be downloaded in batches directly from one mass storage device 34 to another 34' using the associated high speed networks 28, 28' and the high speed transmission pathway 26', with the corresponding database information being transmitted from one mainframe 16 to the other mainframe 16' along the conventional digital transmission pathway 26.

The capture device 32 could be any conventional high speed image capture device, although as discussed herein the capture device 32 and electronic image server 30 are preferably designed, structured, and tuned to match the particular document processing applications for which the high speed network 28 will be utilized.

Referring particularly to FIG. 2, the various components associated with the EIS interface 30 and interconnections between the EIS interface 30 and the host processor are shown.

The host processor includes a communications pathway such as a bus structure 36 which may be of any conventional type defining or defined by the architecture of the 1 host processor. Although an external bus of the type utilized for real time data acquisition and control systems could be utilized, it has proven suitable to design the embodiments discussed herein around the main bus structure 36 of the host processor. It is understood that the bus structure 36 of most personal computers and work stations is hierarchical, including a backbone or system bus, and one or more subsystems such as an instruction exchange bus, a data exchange bus, an arbitration bus to allocate access among competing devices, and an interrupt bus to accept attention request signals from devices. The bus architecture also dictates such features as the motherboard, slot, and card dimensions, grouping and cooling of components, power distribution, and connector configurations and arrangements.

There are currently believed to be twelve accepted full standards for bus structures 36 recognized by the Institute of Electrical and Electronic Engineers (IEEE), and over fourteen proposed standards. There are also many bus standards that have been designed by special interest groups with the cooperation of other organizations to satisfy particular applications, and there are several proprietary bus standards that remain unpublicized. Adapters may also be attached to the system bus to access specialized interface buses such as the Small Computer System Interface (SCSI) for disk drives, scanners, and most other input/output devices, and the

General Purpose Instrumentation Bus (GPIB) for measurement and control devices.

Some of the bus structures contemplated for use with the electronic document transmission, storage, and retrieval system 10 include PCBus (IBM PC class), ISA (IBM AT class), MicroChannel (IBM 32-bit OS/2 class), STD Bus, EISA (Compaq 32-bit), S-100 and SBus (Sun Microsystems), VME (Motorola), open Unibus and QBus (Digital/DEC), and MultiBus and MultiBus II (Intel). It may be readily appreciated that since any component connected to a personal computer bus structure 36 must operate with the host processor and all other peripheral devices, the bus structure 36 of the host processor will dictate many of the design characteristics of the EIS interface 30 and related components. The embodiments subsequently discussed are suited for applications utilizing the ISA bus on IBM compatible computer systems incorporating Intel 386 or 486 processors.

The bus structure 36 of the host processor is preferably equipped for connection to a digital capture device 38 such as a document scanner through a peripheral interface 40, a local area network 12 or peer-to-peer network through a separate network interface 42, an on board or external device capable of image compression 44 using any conventional compression standard such as JPEG, and an output device 46 such as a laser printer of film recorded connected through a serial or parallel output interface 48.

The bus structure 36 may of course be connected to many other types of peripheral devices, however these are the basic devices utilized in the capture, representation, and transmission of electronic image documents over existing networks 12 adapted for interactive document processing.

The operator will have direct command access to both the bus structure 36 and the interface for the EIS interface 30 through a keyboard 50 or other devices such as a mouse, trackball, tablet, or the like, as well as through any input/output and peripheral devices connected to the bus structure 36.

High speed document capture is accomplished using a capture device 52 which provides an analog output signal timed with synchronization and blanking pulses corresponding to individual frames following the normal fast-scan television system standards of 30 frames per second with two interlaced fields or switchable to one progressively scanned field per frame, as modified according to the provisions discussed subsequently herein. The capture device 52 may be any conventional charge-coupled device (CCD) or Vidicon tube type camera, but preferably has a progressive or sequential scan, horizontally and vertically synchronized analog output signal that is switchable between conforming to the fast-scan television system standards producing 30 frames per second and slower vertical sweep rates as required by high resolution modes of operation. An alternative would be to have two cameras to switch between. It may be appreciated that conventional CCD cameras having from less than 512x512 to greater than 2048x2048 pixel arrays will be suitable for many applications employing the modality disclosed herein, and that non-square, larger and small arrays, and linear array cameras or scanners may be equally suited for some applications.

The analog output signal from the capture device is processed through an analog-to-digital converter 54 timed using an input sync-generator 56, with the system being driven by a 124 Mhz clock speed oscillator (crystal) with input to the analog-to-digital converter 54 divided by four to achieve a 31 Mhz conversion clock speed. Since the result-

ant output signal is still a time-division multiplexed signal, a 31 Mhz to 3.9 Mhz demultiplexer 58 may be used to strip the timing and synchronization pulses from the output signal and distribute the remaining image signal to an input memory buffer 60 with a 3.9 Mhz clock rate that is compatible with a broader spectrum of suitable memory devices that are currently commercially available.

It may be appreciated that in the embodiments discussed an 8-bit analog-to-digital converter 54 is used to obtain digital content corresponding to an 8-bit grayscale bitmap, however the continuous analog signal output from the capture device 32 would permit any the use of an analog-to-digital converter 54 that would produce any selected bit level in either grayscale or color depending upon the particular application being considered and informational content desired.

The image input memory 60 is preferably an SRAM semiconductor memory. It should be noted for purposes of this discussion that classification of semiconductor memory usually includes dynamic and static random access memory (DRAM and SRAM), read-only memory (ROM), as well as programmable ROMs (PROM), erasable ROMs (EPROM, EEPROM), nonvolatile RAM (NVRAM), and flash-memory such as charge-coupled devices (CCD) and magnetic bubble memory. Each type of memory may be classified as dynamic or static, read/write or read-only, random or serial access, volatile or nonvolatile, and erasable using magnetic or ultraviolet radiation. Because a CCD sensor chip is properly categorized as a type of serial access semiconductor memory, it is important to distinguish between the bitmap or digital array corresponding to an electronic image formed on the CCD sensor chip of a linear imaging device (LID) such as a line-array scanner or an area imaging device (AID) camera, and the digital array or bitmap associated with the corresponding image once the digital information has been read from the capture device 32, converted and multiplexed for analog television output, interlaced if necessary for video, and subsequently digitized by a frame-grabber or similar device and deposited in active memory which may or may not be user-interactive memory. Consequently, the memory buffer 60 connotes digital semiconductor memory other than the image sensor itself which permits interactive addressing and manipulation of pixels or swapping of information into memory from which pixels may be interactively addressed and manipulate, and which may be configured as an interactive junction with a bus structure 36 or similar interfaces.

The digital array associated with a particular electronic image may be stored to or retrieved from a conventional digital storage device 62 such as a magnetic or optical disk drive, or transmitted to or through any conventional peripheral, through the bus structure 36. It should also be noted that a previously stored electronic image document may be recaptured using the analog capture pathway by substituting an alternate image input source 64 such as a write-once/read-many-times (WORM) optical laser disc player (LDP) connected through a mechanical or electronic switch 66 to the input channel of the analog-to-digital converter 54. A Sony Model LDR-5000A optical disc recorder/player (EIA Standard) has proven suitable.

The switch 66 also provides a suitable junction for linking the input of the analog-to-digital converter 54 to the high speed network 28, and it may therefore be appreciated that an electronic switching device such as a latch controllable by the EIS interface 30 or the mainframe 16 is desired for switching between the capture device 52, laser disc player 64, and high speed network 28 so that electronic images

transmitted on the high speed network 28 will not be obstructed or delayed awaiting manual switching. A manual switch 66 would preferably have a default setting to the high speed network 28, with user-selected switching to the capture device 52 or laser disc player 64 intended only for intermittent operation.

At this point on the image processing pathway, the electronic image document exists as a digital array or bitmap in active memory. The digital array or bitmap may not have the same physical X-Y coordinate array as was present on a CCD sensor chip. In fact, for the applications subsequently discussed, it has proven preferable to divide or section the image into blocks or segments, and to further deposit portions of each block or segment in distinct banks of memory. However, each pixel remains mapped to a specific memory location by bank, row, column, and phase by a matrix or formula in the memory control module 68, and a set digital array or bitmap is preserved and recognized even though it may require calculation to correlate a pixel's location in memory to its position on a raster line or LCD display.

The electronic image document may then be stored on an analog storage device 70 such as an analog WORM optical laser disk recorder (LDR) by dumping the digital array from memory 60 as a conventional 4 Mhz clock speed digital transmission signal and processing that signal through the first step of a sequential frequency converter 72 (up converter) to achieve a 16 Mhz bandwidth, an 8-bit digital-to-analog converter 74, and multiplexer 76 producing a 6 Mhz bandwidth output and the appropriate timing and synchronization pulses expected by the storage device 70.

A switching device 66 may also be utilized to provide a suitable junction for linking the output of the multiplexer 76 to the high speed network 28 for transmission of electronic images, and it may be appreciated that conversion of the signal to the appropriate bandwidth for the high speed transmission pathway may require that the up converter 72 or the multiplexer 76 be instructed by the EIS interface 30 to perform the appropriate signal conversion steps to match the bandwidth of the analog signal to the corresponding transmission pathway.

The 16 Mhz clock rate signal from the first step of the frequency converter 72 may also be processed in a second step to achieve a 124 MHz clock rate signal, converted to analog using a digital-to-analog converter 78, and presented on a conventional raster display 80, preferably such as a high resolution computer monitor. In some applications subsequently discussed, separate input and output memories may be utilized, and the output memory may be read at a speed directly compatible with monitor inputs. The image may also be presented on a conventional LCD display screen 82 or projector, although the necessary signal form may need to be created as required by the LCD display 82.

Since each electronic image document is stored on the storage device 70 in the form of one or more frames (depending upon the operational mode being used and the size of the image), a database 84 is utilized to record an accession number, index, or address for the initial frame and size (or for each of the separate frames) associated with a specific document to permit cataloging and selective retrieval of any document stored on that medium. The database 84 may be located in the mainframe 16 or fileservers, a separate CPU associated with a remote EIS interface 30 device, or be resident in the computer housing the bus structure 36. In many applications involving high speed batch processing on the order of several million

tangible source documents per day (such as checks or similar negotiable instruments processed by a financial institution, for example) each stored image will be identified by a volume number corresponding to the serial number of the specific disc or recording medium, and a frame number. A single frame number will be satisfactory when all batch-processed images are stored using the same quantity of frames for each image, and the combination of the volume and frame numbers may be treated together as a single index, address, or accession number. The database 84 may be an already-existing database used for processing transactions involving the source documents to which one or two fields are appended corresponding to the volume and frame numbers for the stored image, or may be a separate database 84 which provides a "look-up table" relating an identifying characteristic of the source document (such as the computer-readable MICR code on negotiable instruments or checks) with the corresponding serial and frame numbers of the stored image.

Because the speeds at which tangible documents may be captured exceeds the rate at which an operator can manually position the document, it is anticipated that automatic document feeders (not shown) will be utilized particularly with paper, film, and microfiche documents. Consequently, an infed control interface 86 must be connected to the EIS interface 30 so that the EIS interface 30 can receive ready-state acknowledgements from the automatic document feeder and step or advance commands can be sent to the automatic document feeder, and a corresponding indexing or accession control interface 88 would be connected between the EIS interface 30 and the storage device 70 to correlate the volume serial numbers and frame numbers with the appropriate images that are stored, and provide sequencing and control signals to the storage device 70. A manual infed control interface 86 would be used when documents are positioned manually by the operator.

A separate storage device 70' such as a second analog WORM optical laser disc recorder (LDR) may be utilized at a remote location for batch downloading or backup of images stored on the first storage device 70, with the corresponding indexing or cataloging information from the database 84 being supplied to a separate database that is connected to the second storage device 70'.

#### Source Document Orientation

The embodiments of the system 10 discussed herein have been specifically tailored to optimize the processes of capturing and storing electronic images from tangible source documents that are especially common to many financial and business transactions conducted in the United States, or for specialized applications that are considered document intensive. Those tangible source documents include paper documents, as well as documents previously fixed in mediums such as microfilm and microfiche. The following discussion of several paper-based source documents provides a sufficient basis for describing the particular modes of operation associated with the preferred embodiments.

Referring particularly to FIGS. 3-7, several of these source documents are shown. FIG. 3 is a diagrammatic depiction of a basic US letter-size page 90 generally approximating the horizontal and vertical dimensions or aspect ratio of a standard 8½"x11" sheet of paper (and similarly approximating an A-4 size page) in a vertical orientation. FIG. 4 is a diagrammatic depiction of the same page 90 shown in FIG. 3 in a horizontal orientation. The

page 90 is divided along an imaginary vertical centerline 92 into two half-page segments 94, 96 having dimensions of approximately 8½"x5½" each. Two blocks 98 each representing the front or back face of a standard check (negotiable instrument) having dimensions of approximately 2¾"x6" are superimposed on the right half 96 of the page 90. FIG. 5 is a diagrammatic depiction of the page 90 of FIG. 3 showing the normal orientation of type written lines 100 of text on that page 90.

FIG. 6 is a diagrammatic depiction of a drawing sheet 102 having dimensions greater than twice a basic page 90, with horizontal and vertical phantom lines 104 showing the drawing sheet 102 divided into four equal quadrants or segments along both horizontal and vertical seam lines.

FIG. 7 is a diagrammatic depiction of a section of a large-scale document 106 such as a geophysical survey. The section 106 has dimensions on the order of 36"x48" and may correspond to about 4 miles of survey data with 50 samples per mile along the horizontal axis, with the entire document encompassing 70 or more miles of survey data and having a total length of many feet.

#### Capture Device Arrays

Referring particularly to FIGS. 9-13, it may be seen that the processes of capturing and storing electronic images from the types of source documents 90, 94, 98, 102, 106 discussed above may be accomplished in several manners depending upon the corresponding size of the source document 90, 94, 98, 102, 106 and the degree of informational content desired to be captured and stored.

FIG. 9 is a diagrammatic depiction of a single capture device 32 such as an area array CCD or Vidicon tube type camera disposed a predetermined distance or height above a source document such as a basic page 90 oriented horizontally, with a phantom line 92 identifying an imaginary centerline 92 dividing the basic page 90 into two half-page (one frame) segments A, B. FIG. 10 is a diagrammatic depiction of a pair of capture devices 32 disposed above a source document 108 such as a basic page 90 oriented vertically or two-page drawing sheet 102 oriented horizontally, with a phantom line 110 identifying an imaginary centerline or seam line dividing the source document 108 into two segments A, B (of two frames each).

FIG. 11 is a diagrammatic depiction of a pair of capture devices 32 disposed above and below a source document 112 such as a check, with one capture device 32 capturing a segment A of the resulting half-page electronic image corresponding to the front face of the check or source document 112, and the other capture device 32 capturing a segment of the resulting half-page (one frame) electronic image corresponding to the opposing face (not shown).

FIG. 12 is a diagrammatic depiction of a plurality of capture devices 32 disposed in a two dimensional (vertical and horizontal) array above a large-size source document 114, with each capture device 32 capturing a segment A-F of the resulting (multi-frame) electronic image which adjoin one another along both vertical and horizontal seam lines 116.

FIG. 13 is a diagrammatic depiction of a capture devices 32 of the line array type disposed above a large-size source document 114, with the capture device 32 sequentially capturing a segment A-C of the resulting (multi-frame) electronic image which adjoin one another along parallel seam lines 116.

### "Mode A" and "Mode B" Operation

Different operational modes for the electronic document transmission, storage, and retrieval system 10 may be designed to achieve particular goals or meet certain requirements, limitations, or constraints associated with the document or the system 10. Two modes termed "Mode A" and "Mode B" are particularly designed around document processing involving tangible source documents such as a US letter-size page 90 or smaller, 8-bit grayscale informational content, and wherein "normal" system resolution is suitable. It should be noted that "normal" resolution refers to an arbitrary resolution as defined by the particular system mode and the application of that mode to a specific operating environment. In other applications, normal resolution might refer to fractions or multiples of the "normal" resolution as defined by Modes A and B. It should also be noted that bypassing the bandwidth filters or bandwidth limiters in conventional analog WORM-type LDRs 70 to operate with the "extended" bandwidth and additional lines of image content results in an approximately 30% increase in the effective resolution or storage capacity of the storage medium, without affecting its storage density. These modifications are possible because the LDRs were designed around color broadcast specifications which differ from the optimal performance arrangement for document legibility.

Modes A and B are particularly adapted for a 1-frame 30 hz or 2-frame near 14 hz (i.e., less than 15 hz or approximately 70 millisecond) record time, capturing informational content in 8-bit grayscale from a horizontally-oriented  $8\frac{1}{2} \times 11$ " tangible document. This orientation is believed preferable for documents containing typewritten characters, because readability of a typewritten characters requires more resolution in the "horizontal" direction than in the "vertical" direction. It may be noted that the frame rate of Mode B is deliberately not one half the frame rate of Mode A, since tube type cameras require a minimum vertical retrace time and it the necessity of achieving 1000 plus useable lines.

Mode A refers to capturing and storing one half-page image of a US letter-size source document as a single frame, whereas Mode B refers to capturing and storing the full-page image of a US letter-size source document as two frames. Although Mode A can be considered a system bypass that does not require complex seaming to retrieve and reconstitute the electronic image, the same control signals including the pilot signal and calibration pulses as discussed below are applied to documents processed in Mode A, since those control signals are also useful for maintaining electronic image integrity throughout batch-processed source documents and between different operating environments or platforms.

Normal fast-scan television provides 30 frames per second with 525 horizontal lines swept at a 60 hz vertical rate and scanned at a 15,734 hz horizontal rate with a 4:3 aspect ratio, with two interlaced scanned fields per frame. The image is sent as an amplitude-modulated (AM) signal, while audio is frequency-modulated (FM). European television (PAL standard) utilizes 625 lines, however the aspect ration remains approximately 4:3.

In Modes A and B, the capture device 52 produces a camera image of 1114 progressive (non-interlaced) lines refreshed every 70 milliseconds. When recorded using the method discussed herein on a conventional laser disc recorder 70 which expects a 525 line format, this will correspond to two images of 500 lines each with the excess lines accommodating the control signals as described. The lines of the image are therefore being scanned at 14 hz in the

vertical direction (approximately one quarter the normal sweep rate) and scanned horizontally at 15 khz. A Dage-81 camera manufactured by Dage-MTI Incorporated of Michigan City, Indiana, provides a suitable capture device capable of being driven at the 14 hz vertical scan rate. It should be noted that more accumulated charge may be discharged as useable signal by lengthening the time during which a document is being scanned using a capture device in which the accumulated charge on the target is time-dependent on incident source light such as a Vidicon tube type camera. This results in enhanced image capture capabilities by effectively increasing the number of "pixels" that are contained in each scan line, although the number of scan lines that may be treated as a "frame" is reduced proportionately because each scan line is longer in duration, and the image therefore occupies more frames. Tube type cameras are therefore advantageous in applications where a CCD type camera would not provide sufficient resolution or fast enough speeds given other practical considerations, or where the fixed pixel resolutions of current CCD cameras are otherwise inadequate. While this scan rate reduction would normally produce a visible flicker if the images being captured were viewed directly from the camera output, the scan rate reduction is compensated for by subsequent frequency conversions that are interposed immediately prior to any display, projection, or other visible presentation operation. Progressive scanning is preferred since it provides a Kell factor as high as 0.9 over the traditional 0.7 value for interlaced scanning, thus producing advantages in acquisition and display, and since interlaced scanning results in greater information loss caused by inaccuracies in tracking adjacent sweeps and the time factor for decay occurring between interlaced sweeps.

The output signal from the capture device 52 is digitized at 31 Mhz and stored in a  $1024 \times 2048 \times 8$  input memory buffer 60. The memory buffer 60 is preferably two banks of eight 128 kbit SRAMS with an 8 column by 9 row architecture and a corresponding video input latch. The output signal is multiplexed into 8 segments which places 256 pixels in each row of memory. Any pixel can be identified by bank, row, column, and phase. The configuration of the memory buffer 60 in Modes A and B can be made more efficient, however this configuration permits interchangeable application to other modes and is simple to construct.

It may be appreciated from this discussion that the true potential of the  $1024 \times 2048$  memory buffer 60 configuration is not necessarily being fully utilized in the embodiments which employ currently available state-of-the-art LDRs, however use of this configuration of the memory buffer 60 provides a suitable margin to permit the use of these embodiments with improved LDRs currently under development with a minimum of adaption. As noted above, there is also a rationale for having separate input (capture) and output (display or storage) memory buffers, with each separate memory buffer being optimized for its intended purpose. Since the output from the input memory to the storage device is analog and the input to the output memory is also analog for retrieval, commonality of the digital characteristics of input and output memory is unnecessary. Separate memory buffers also facilitate simplicity in designing the interface modules to have retrieve-only capabilities, thus reducing the cost and complexity per node for network systems having a centralized high speed capture and storage process.

The EIS interface 30 circuitry may be divided into three groups. The circuitry associated with the memory buffer 60 is one group. The non-124 Mhz portion can be fabricated on

35

a first integrated circuit board using transistor-transistor logic (TTL), whereas the 124 Mhz portion (including the digital-to-analog converter 74) can be fabricated as a third group using emitter-coupled logic (ECL) on a separate integrated circuit board. The multiplexer 76 can be fabricated with the non-124 Mhz portion of the circuitry, or a remote multiplexer 76 may be connected to the non-124 Mhz portion of the circuitry using sixty-four sets of twisted-pair wires.

The resultant analog signal associated with the electronic image may be stored on the storage device 70 in video format. A 525 line 60 hz vertical scan rate WORM optical laser disc recorder (LDR) is conventional, and a Sony Model LVR-5000A (EIA Standard) laser videodisc recorder/player has proven suitable.

Storing a 1024 line image in two standard frames produces 26 excess lines, for which the LDR would "see" approximately one fourth the normal number of horizontal synchronization pulse durations during the vertical interval instead of the 21 expected, and this could affect the LDR's servo lock or otherwise result in inoperability of the LDR. There are two immediately available options for resolving this situation. First, the LDR could be adapted to expect the narrow fine vertical intervals. However, since the use of unmodified standard LDRs is desired, the preferred embodiment utilizes a standard vertical interval signal with a 1000 line video image, which necessarily results in an image that is 24 lines short of binary "pure."

It may be readily appreciated that a system utilizing a 625 line 50 hz vertical scan rate LDR compatible with the PAL (European) television and video recording standard may be preferred in situations where PAL-compatible LDRs are more readily accessible or in applications where the nature of the tangible source document makes 625 line capture the optimal alternative. In some applications, the increase in horizontal bandwidth from 4.2 Mhz to 6 Mhz may be compelling. Such a conversion will either decrease the number of images storable on a conventional optical disc by 20% (from approximately 43,000 to 36,000) or increase the recording time by about 20% per frame. In such an event, the sync generator 56 can be modified to operate corresponding to the 625 line 50 hz vertical scan PAL standard. However, it should also be remembered that the sync generator controls the drive frequency of the capture device 32 and supplies the timing waveforms for the input memory buffer 60. Consequently, selection of the appropriate capture device 32 and optimization of memory characteristics may be required.

A 525 line 60 hz vertical scan rate standard sync generator provides a composite synchronization train. After the first vertical synchronization pulse, the first 250 lines are read to the memory buffer 60. The read pauses to allow insertion of the second vertical interval, and then lines 251-500 are read. The read pauses, and the vertical interval between the first and second frames is inserted, and the second frame is read in the same manner as the first frame. The additional 50 line times are consumed by the vertical blanking intervals. At the same time the read starts for the beginning of the first frame, the LDR begins recording the output from the digital-to-analog converter 74. The image can be retrieved by reversing the storage process and reconstituting the two frames as a single electronic image document.

It may be appreciated that the tangible  $8\frac{1}{2} \times 11$  inch original image is divided into four blocks or segments to be stored on two separate frames with two blocks per frame. These blocks or segments can be groups of sequentially adjacent

36

lines, fractions or segments of adjacent lines, alternating or selected interlaced lines, or any other desired portion of the image. An electronic image having a digital array composed of X lines of pixels would result in two frames each having X/2 lines of information, and the raster corresponding to those two frames would be composed of N horizontal scanning lines at a predetermined vertical scan rate plus a maximum time interval dictated by the remaining or excess 2N-X horizontal scanning lines and the predetermined vertical scan rate which is utilized to add the raster synchronization to the analog signal for each frame.

The retrieved electronic image can be processed in two ways. It can be mapped at 3.9 Mhz onto the bus structure 36 of the host processor and swapped to the video display RAM on the computer's motherboard or on a video/graphics interface card, stored in digital form on magnetic medium, transmitted over conventional transmission pathways, or downloaded to an output device such as a printer. Alternately, for high resolution display the signal may be up converted by the frequency converter 72 to correspond to the appropriate dot clock equivalent for a high resolution gray scale or color monitor.

#### "Mode C" Operation

Modes A and B operation are considered biased in the horizontal direction. Mode C operation is effective where either "square" resolution or a vertical bias are desired. Mode C operation may be considered "high" resolution compared with Modes A or B, and provides twice the "normal" resolution of those modes.

The acquisition of the signal in Mode C is the same as in Modes A and B, however the image is dissected into 8 blocks and recorded as 4 frames. The resolution of the tangible source document image (using the  $8\frac{1}{2} \times 11$  inch example) becomes a 1024M X-Y pixel product (1248 pixels in the 11 inch direction and 1000 pixels in the 8.5 inch direction) while only requiring twice the capture and storage time.

There are three alternatives for operation in Mode C. In each case the horizontal line is split in half. The read from the memory buffer 60 is performed at half the write speed (using a 7.5 khz horizontal sweep rate controlled by the read clock) to effectively expand or stretch the 12 Mhz bandwidth signal into a 6 Mhz bandwidth signal, which causes the half of the horizontal line segment to expand from 32 microseconds to 64 microseconds in duration, which appears normal to an LDR expecting a 6 Mhz analog signal with 64 microsecond line duration.

Each of the three alternatives to Mode C may be understood by considering a grid overlying a 1000 line raster image in which each line has a left and a right half.

In the first alternative (Mode C1 in FIG. 15) the grid has one vertical column of eight horizontal rows. Each row of the grid therefore corresponds to one field of a frame. Each segment is therefore one vertical row consisting of 125 left halves alternating with 125 right halves of each line. The segments are stored in ascending order, S1 through S8, so that frame F1 contains segments S1 and S2 as fields 1 and 2, frame F2 contains segments S3 and S4 as fields 1 and 2, and so forth. This mode is the least expensive since it only requires a line delay, and has the advantage of providing a security function by "scrambling" the appearance of the stored images in the event an unauthorized individual attempts to access the medium using an incompatible EIS system or view the image on an interlaced monitor.

In the second alternative (Mode C2 in FIG. 16) the grid has two vertical columns of four horizontal rows. Each

segment therefore consists of either 125 left halves or 125 right halves of the lines. The segments are again stored in ascending order, S1 through S8, so that frame F1 contains segments S1 and S2 as fields 1 and 2, frame F2 contains segments S3 and S4 as fields 1 and 2, and so forth. This mode is the more expensive since it requires a field delay, and also has the advantage of providing a security function by scrambling the appearance of the stored images. The security feature of Mode C2 would produce significantly more flicker because of the difference between the horizontal sweep and vertical scan rates.

In the third and fourth alternatives (Mode C3 and Mode C4 in FIGS. 17 and 18, respectively) the grid is the same as for Mode C2, and each segment therefore consists of either 125 left halves or 125 right halves of the lines. In Mode C3 and Mode C4, however, the segments are not stored in ascending order. Instead, in Mode C3 the left halves of the lines corresponding to two segments S1 and S2 are stored as the first and second fields in frame F1, and the right halves of the lines corresponding to two segments S5 and S6 are stored as the first and second fields in frame F2. The remaining segments S3, S4, S7, S8 are similarly stored as frames F3 and F4.

In Mode C4, the left halves of the lines corresponding to two segments S1 and S2 are stored as the first and second fields in frame F1, and the left halves of the remaining lines corresponding to segments S3 and S4 are stored as the first and second fields in frame F2. The right halves of the lines corresponding to segments S5 and S6 are stored as frame F3, and the right halves of the lines corresponding to segments S7 and S8 are stored as frame F4.

Modes C3 and C4 have the advantages of allowing the direct display of a quadrant (two adjacent segments) of the image on a conventional sequential scan (non-interlaced) monitor, and in the event of discrepancies in or uncertainty about the information in the overlapping regions between quadrants, the adjacent quadrants may be viewed separately for comparison. Mode C3 presents the disadvantage of requiring a one frame delay in order to seam the quadrants properly, and Mode C4 requires a two frame delay. Mode C3 is therefore more expensive than Mode C2, and Mode C4 more expensive than Mode C3. Mode C4 may provide an advantage in simultaneously viewing two vertically adjacent quadrants which represent a strip of an electronic image or tangible document which has a extremely long length equivalent to a continuous roll of paper.

In the second alternative (Mode C2) is to place each half of the stretched line sequentially one after another. The first vertical block or field would therefore have 256 lines of active video comprising 128 first halves alternating with 128 second halves. This continues for 8 blocks or fields, at which point 4 frames have been stored. To retrieve an image stored in this mode, the added sync signals are removed and the second halves are delayed and joined to form continuous lines having a 12 Mhz bandwidth.

When an image is retrieved in Mode B, the blocks or segments corresponding to adjacent lines will seam together naturally because the separation between blocks is equivalent to a normal horizontal or vertical synchronization pulse. In the case of line segments, the top quadrants will seam together with their adjacent bottom quadrants, however a more complex seaming technique is preferably utilized between horizontally adjacent quadrants.

One technique for seaming horizontally adjacent quadrants is overscanning, in which the first halves each line segment extend 2% beyond the midpoint of the individual

line segment to include a leading portion of the second half of the same line (in another segment and possibly in another frame), and the second halves begin 2% before the midpoint to include a trailing portion of the first half of the same line (also in another segment and possibly in another frame). The overlapping or redundant information may be compared in dc or ac content, time, and amplitude to match the seam, with servo loops being employed to provide the necessary degree of precision required in the manner discussed below. Higher frequency content and increasing grayscale depth will dictate the type of servo loops used.

#### "Mode N" Operation

Mode N refers generally to any operating mode which employs one or more capture devices 32 to capture, dissect, and seam together or reconstitute an electronic image corresponding to a source document larger than a basic US letter-size page at "normal" resolution, or to a source document of any size that is captured and stored as more than four frames and more than eight segments. It may be readily appreciated from this discussion that it various embodiments may be designed around the use of a single capture device 32 with the source document being moved or advanced to present different regions that would correspond to different segments of a larger electronic image or document, or that several capture devices 32 could be utilized as shown in FIGS. 10 or 12, either with or without movement or advancing the source document. A line-array type capture device 32 as shown in FIG. 13 may be more practical for some continuous feed applications in which the documents have a fixed width but variable length.

In some applications, such as the example provided above relating to negotiable instruments or checks, an array of two capture devices 32 such as shown in FIG. 11 are utilized to capture an images relating to the front and back faces of the source document, and those two images are subsequently placed together as a single half-page segment or frame. It may also be appreciated that several segments from several different capture devices 32 can be seamed together without regard to whether each of the segments represents a portion of a larger source document, thus allowing new documents to be created by assembling segments of other images together as though reconstituting and original electronic image.

#### Seaming and Image Control Signals

The raster synchronization discussed above that is added to the analog signal in order to define frames corresponding to the selected storage medium is a conventional format that may be recognized by commercial and consumer video equipment, however such a format is only a "coarse" timing reference compared to the degree of accuracy or precision necessary to dissect and reconstitute electronic image documents and still maintain the requisite level of qualitative integrity for the informational content described.

The vertical and horizontal synchronization pulses embedded in standard RS-170 formatted signal (used with standard television display and recording) are unsuitable for seaming adjacent segments of an electronic image. Four types of image control signals are therefore utilized with the embodiments of the system 10 disclosed herein.

The first two sets of signals are the horizontal and vertical raster synchronization pulses (and corresponding blanking intervals) corresponding to the camera output and frame formats discussed above. Referring to FIG. 14a, a 50 micro-

second sampling of a 12 Mhz information signal 118 is shown bracketed by a pair of 13.5 microsecond blanking intervals 120 and a pair of 5 microsecond negative-going synchronization pulses 122. The line thus has a duration of approximately 63.5 microseconds. When digitized or quantized at a 31 Mhz conversion clock rate this produces a 50 microsecond digital signal as shown in FIG. 14b, which may be segmented and stretched into two halves which are each 51.5 microseconds in length by reading from the memory buffer 60 at 15.5 Mhz, thus producing two spaced-apart signals as shown in FIG. 14c and an intervening dead zone 124. The two spaced-apart signals are converted to analog signals as shown in FIG. 14d, with a portion constituting approximately 2% of each half line situated directly adjacent to the dead zone 124 accounting for the overlap 126 created by and utilized for overscanning.

The third type of image control signal are calibration pulses 128 embedded as a part of the video signal just before and after the active video portions of the composite signal. The calibration pulses are positive-going pulses inserted immediately before and after the synchronization pulses for each line (or line segment). Each calibration pulse preferably has a width (duration) on the order of 750 nanoseconds and a height (amplitude) of 0.35 volts DC when a 0.714 volt active video signal is utilized. The calibration pulses ensure matching of the DC levels and amplitudes of adjacent segments, incremental gain calibration between adjacent lines (or line segments) of the image and against a set level for each line throughout the subject image and prior or subsequent images, and also provide an intermediate timing reference that has a resolution greater than the conventional synchronization signals but less than the pilot signal.

Gain calibration is accomplished by subtracting black from gray, with true black being clamped at zero. If a calibration pulse is read at 0.4 V, for example, a servo correction brings that line down to 0.35 V to provide an accurate gray level for each "pixel" within the line. The calibration pulse at the end of any line may be directly compared with the calibration pulse at the beginning of any adjacent line along the corresponding seam. Any divergence from the normal calibration amplitude results in shifting the level of the corresponding line so that each line exactly matches the adjacent lines and is set relative to a fixed base level, and the servo loops used for the calibration adjustments are therefore completely independent of informational content in the image.

It may be appreciated that the calibration pulses 128 provide reference to a predetermined "absolute gray" that also corresponds to a similar value for monitors and other devices, with the focus of the embodiments discussed being on a predetermined area of interest or range of grayscale levels that are commonly encountered in tangible source documents of the type used for business records, financial transactions, and so forth. The amplitude of the calibration pulses 128 may be set at any desired value to provide control over the location of the examining area within the entire grayscale range by shifting the median grayscale of the image upward or downwardly compared to an "absolute gray" value expected by the monitor or other device, while still permitting "on the fly" correction around the calibration pulses 128 for line-to-line and segment-to-segment precision.

A conventional negative-going synchronization pulse 130 is then inserted between the calibration pulses 128, as shown in FIG. 14f.

The fourth type of image control signal is a low level (CW) pilot signal 132 or carrier tone that extends continu-

ously throughout the stored analog signal, is synchronous with the calibration pulses 128, and is harmonically related to the clock speed. The pilot signal 132 must be high enough frequency to permit harmonization with all other synchronizing or timing signals and pulses, and therefore compatible with all divisors of the main system clock rate, and is preferably at or near one half of the write clock frequency. The pilot signal 132 should also be a frequency just above the predetermined bandwidth of the storage medium, but sufficiently distinguishable that a bandpass filter can be used to strip the pilot signal 132 without clipping active video content 118 from the signal. Since the operational spectrum for conventional laser disc medium falls off rapidly above the 6 Mhz bandwidth, a low level pilot signal 132 on the order of 6.89 Mhz has proven suitable for use with conventional LDRs, and may be filtered out prior to display of the electronic image to prevent visual disruption of the displayed image. The pilot signal 132 therefore provides many times more vernier than the standard synchronization pulses for time-based corrections in positioning and aligning adjacent segments of the image when seaming those segments together along horizontal seam lines (parallel with raster lines) or vertical seam lines (perpendicular to raster lines). The pilot signal 132 is also necessary to provide a time-based correction capability to compensate for timing errors in conventional LDRs, and the use of the recovered pilot signal 128 which is retrieved as a part of the stored signal to generate the write clock as data is written to the memory buffer 60 allows the use of this modality in connection with other storage and communications systems that produce time-based errors or which do not regulate absolute timing relative to an external reference or source.

When the pilot signal 132, calibration pulses 128, and synchronization pulses 130 of FIGS. 14e-14g are summed with the active signal portion of FIG. 14d, a composite analog waveform 134 results. As shown in FIG. 14h, that waveform generally follows and oscillates about the summed control signals 128, 130, 132.

To reconstitute the image, the composite signal 134 of FIG. 14h is retrieved from the storage medium, and a square wave from the write clock at 15.5 Mhz (derived from a recovered 6.89 Mhz pilot signal) is used to quantize or digitize the signal and write the signal to the memory buffer 60 in the same form as shown in FIG. 14c. The "left hand" segment of the digitized signal is delayed to overlap with the "right hand" segment, and the two segments are seamed together as shown in FIG. 14b, and the digital array is converted to an analog signal with raster synchronization and blanking intervals added as shown in FIG. 14a, thus permitting the storage, transmission, or display of the electronic image. In the four versions of Mode C operation, the delay will be a line, field, or frame delay depending upon the mode.

For an 8-bit grayscale image, the processed signal does not correspond to a 1:1 quantization that would produce a true 256 level grayscale range for the electronic image. Instead, it has proven suitable in the embodiments discussed to incorporate an approximately 20% margin to provide overhead for calibration adjustments and offset. As such, the effective grayscale range would be on the order of 200 to 210 levels, which accommodates the vast majority of interactive document processing applications including high resolution medical imaging for X-rays and MRIs, and is state-of-the-art for commercially available LDRs. Applications requiring enhanced grayscale differentiation may result in sacrificing a portion of the margin, or increasing the bit depth of the system.

## Transmission Pathways

The transmission of an electronic image as an analog signal provides significant improvement in transmission rates compared with even the fastest conventional digital transmission by modem or facsimile or over network lines. However, it may be appreciated that the segmenting of the electronic image in the memory buffer 60 and conversion to a resultant analog signal of appropriate bandwidth permits additional increases in transmission rates over conventional transmission pathways having limited bandwidths, as well as high speed transmission pathways having larger bandwidths.

It is important to note when considering conversion of an analog signal for transmission that the parameters for transmission over a digital transmission pathway are traditionally specified in Mhz clock speeds. According to the Nyquist theorem, a perfect system free of filtering imperfections would permit the transmission of an analog signal having a frequency (R) referred to as the Nyquist frequency at no less than twice (2R) the clock rate which is referred to as the sampling or Nyquist rate. In the preferred embodiments discussed above, the Nyquist frequency of the analog signal from the capture device 32 and from the initial digital-to-analog conversion is on the order of 12 Mhz, whereas the sampling rate uses a 31 Mhz clock rate. Consequently, this produces a Nyquist ratio of approximately 2.6 between the clock rate and analog Nyquist frequency. This conversion ratio also provides an estimate for the effective analog bandwidth that may be allotted to a specific digital pathway. As such, if one were transmitting a 12 Mhz analog signal over a 4 Mhz clock rate transmission pathway in a perfect system (where the Nyquist ratio is 2 and the maximum bandwidth is 2), it would be necessary to either slow the analog signal down by a factor of 6 times and transmit continuously, or divide the signal into six components and transmit in parallel.

Utilizing the system 10 of this invention, it is possible to accomplish either of these processes in a way that is optimal for the particular application and transmission pathway. The actual bandwidth of the analog signal may be reduced to a lower bandwidth that corresponds to a digital clock rate at or below the maximum transmission clock rate of the digital transmission pathway, and then transmitted at that speed (which would be slower than if transmitted at the original bandwidth over a line having the same effective bandwidth or clock speed.) Alternately, the electronic image may be divided into a plurality of segments each having an effective bandwidth corresponding to the maximum transmission clock rate of a particular digital pathway, and then transmitted in parallel and reconstituted. In a transmission pathway having more than one parallel lines or channels (such as a conventional ISDN transmission pathway comprising three parallel channels of which two are allocated for digital transmission), the segments of the electronic image in the memory buffer 60 can be read in parallel and transmitted simultaneously on the separate digital channels or lines. Again, the bandwidth of the analog signal for each channel is optimized according to the operating parameters of that channel. Subsequent frequency conversion of the type obtained using a line multiplexer may also be applied in applications such as fiber-optic transmission, permitting a multiplicity of segments of an electronic image to be transmitted simultaneously. The analog signal may also be divided accordingly and transmitted over an analog transmission pathway at the maximum bandwidth permitted.

In their simplest embodiments, the high speed network 28 and transmission pathway 26 are broadband cables capable

of transmitting television-type signals, utilizing a conventional handshake recognition and latch to lock out non-requesting nodes from the communication procedure once a request is received by the mainframe or controller and an instruction to retrieve and transmit an image is received and completed by the electronic image server 30 and mass storage device 34. Image requests, sequencing or cataloging information, and ready-state or other control signals may be transmitted over a conventional digital network such as a Novell or Ethernet system.

Because the transmission bandwidths and number of available channels or pathways are fixed for conventional communications networks, the system 10 may be optimized to permit transmission on these existing communication lines, including satellite and microwave transmission, multichannel RF television, as well as transmission over home-broadcast cable television systems. The transmission operation is also completely compatible with and transparent to any conventional communications-related security technology, such as those employing modulated line scrambling devices and signal encryption algorithms.

It may be readily appreciated that the various embodiments and modes of operation discussed above constitute only representative examples of the optimization of the disclosed method to: (1) a specific size and type of source document; (2) a particular operating environment; (3) a predetermined level of informational content or resolution for the electronic image; and (4) the utilization of selected devices and equipment for initial image capture, CPU bus and interface, storage medium, and transmission pathways based upon existing commercial availability and cost.

It is understood that the systems for practicing the disclosed method may be optimized according to many different parameters beyond informational content and processing time, however it is expected that these two parameters will remain most significant for practical commercial applications. It is further understood that many factors will affect the design of alternate modes of operation or further embodiments of the system, such as: the selection of alternate devices or equipment; the continuing refinement and introduction of new capture, storage, and transmission technologies; the operational guidelines imposed for accomplishing specified tasks for certain applications; the nature and limitations of the operating environment; the modifications or adaptations involved in applying this method to electronic content; the transition between different television, video, and communications standards in this and other countries; as well as the peculiar emphasis that individuals or users may be placed on etherial considerations such as cost, complexity, security, or other aesthetic values affecting basic design features.

While the preferred embodiments of the above electronic document transmission, storage, and retrieval system 10 has been described in detail with reference to the attached drawing Figures, it is understood that various changes and adaptations may be made in the electronic document transmission, storage, and retrieval system 10 without departing from the spirit and scope of the appended claims.

What is claimed is:

1. A method for the storage and retrieval of a two-dimensional electronic image expressible as a digital array or bitmap for the purpose of interactive document processing, said method being utilized in association with a storage device having a storage medium, said method comprising the steps of:

writing at least a portion of the digital array or bitmap corresponding to the two-dimensional electronic image to an output memory;

reading said digital content from said output memory;  
 converting said digital content to a first analog signal as  
 said digital content is read from said output memory;  
 adding a plurality of image control signals to said first  
 analog signal to define a format, said format defining a  
 frame;  
 selectively transmitting said first analog signal over the  
 transmission pathway to the remote location or storing  
 said first analog signal as at least one said frame on the  
 storage device, the storage device being capable of  
 receiving and storing said first analog signal on the  
 storage medium along with a multiplicity of like  
 frames, the storage device further being capable of  
 randomly accessing and selectively retrieving and out-  
 putting a separate analog signal corresponding to said  
 first analog signal stored as at least one said frame from  
 among said multiplicity of like frames, said separate  
 analog signal including said plurality of image control  
 signals;  
 retrieving said separate analog signal including said plu-  
 rality of image control signals from the storage device;  
 stripping said plurality of image control signals from said  
 separate analog signal retrieved from the storage  
 device;  
 converting said separate analog signal retrieved from the  
 storage device to a digital signal;  
 writing said digital signal to an input memory such that  
 the digital array or bitmap corresponding to the two-  
 dimensional electronic image is reconstituted in said  
 input memory,  
 whereby the electronic image may subsequently be uti-  
 lized in one or more document processing operations  
 including the transformation, presentation, representa-  
 tion, transmission, or storage and retrieval of the elec-  
 tronic image,  
 wherein the plurality of image control signals includes  
 a plurality of raster synchronization pulses and cor-  
 responding blanking intervals, at least one pilot  
 signal, and at least one set of calibration pulses.

2. The method of claim 1 wherein the storage medium has  
 a predetermined bandwidth and wherein the frequency of the  
 pilot signal is at or slightly above said predetermined  
 bandwidth of the storage medium.

3. The method of claim 2 wherein the predetermined  
 bandwidth of the storage medium is on the order of 6 Mhz  
 and the pilot signal is on the order of 7 Mhz.

4. The method of claim 1 wherein the pilot signal is a  
 continuous signal which extends throughout the separate  
 analog signal.

5. The method of claim 1 wherein the format defining the  
 frame is a raster composed of N scanning lines at a prede-  
 termined scan rate, each of said scanning lines having an  
 active portion corresponding to an informational content of  
 the two-dimensional electronic image, and further wherein  
 the at least one set of calibration pulses include a first  
 calibration pulse inserted before the active portion of each of  
 the scanning lines and a second calibration pulse inserted  
 after the active portion of each of the scanning lines.

6. The method of claim 5 wherein the first calibration  
 pulse and the second calibration pulse each have a duration  
 and an amplitude, such that said amplitude of the first  
 calibration pulse of a first scanning line may be compared  
 with either said amplitude of the first calibration pulse or  
 said amplitude of the second calibration pulse of a like  
 scanning line or with a fixed amplitude.

7. The method of claim 6 wherein each of the scanning  
 lines has a gain associated therewith, and wherein said gain

of a selected scanning line may be adjusted in response to a  
 comparison between the first calibration pulse of said  
 selected scanning line and the second calibration pulse of a  
 like scanning line or with a fixed amplitude.

8. The method of claim 6 wherein the duration of the first  
 calibration pulse and the second calibration pulse is on the  
 order of 750 nanoseconds.

9. The method of claim 6 wherein the amplitude of the  
 first calibration pulse and the second calibration pulse is on  
 the order of 0.35 volts.

10. The method of claim 1 wherein the at least one set of  
 calibration pulses provides an intermediate timing reference  
 that has a resolution greater than the plurality of raster  
 synchronization pulses and corresponding blanking intervals  
 and less than the pilot signal.

11. The method of claim 1 wherein the digital array or  
 bitmap corresponding to the two-dimensional electronic  
 image is composed of X lines, and wherein the format  
 defining the frame is a raster composed of N scanning lines  
 at a predetermined scan rate.

12. The method of claim 11 wherein the frame is com-  
 posed of X/2 lines, and a time interval dictated by a quantity  
 of 2N-X excess scanning lines and the predetermined scan  
 rate is utilized for containing the plurality of image control  
 signals.

13. The method of claim 12 wherein X is on the order of  
 1024, N is on the order of 525, and the predetermined scan  
 rate is on the order of 60 hz.

14. The method of claim 1 wherein the method further  
 comprises the step of:

sectioning the digital array or bitmap corresponding to the  
 two-dimensional electronic image into a plurality of  
 segments prior to or during the step of writing the  
 digital content to the output memory.

15. The method of claim 14 wherein the number of the  
 frames is one, and the number of the plurality of segments  
 is two.

16. The method of claim 14 wherein the number of the  
 frames is two, and the number of the plurality of segments  
 is four.

17. The method of claim 16 wherein the digital array or  
 bitmap corresponding to the two-dimensional electronic  
 image is composed of X lines, and wherein each segment  
 corresponds to a block of X/4 adjacent lines.

18. The method of claim 17 wherein the plurality of  
 segments includes a first segment, a second segment, a third  
 segment, and a fourth segment, said first segment being  
 adjacent to said second segment and said third segment  
 being adjacent to said fourth segment, and wherein the  
 number of the frames includes a first frame and a second  
 frame, said first segment and said second segment being  
 stored in said first frame and said third segment and said  
 fourth segment being stored in said second frame.

19. The method of claim 14 wherein the portion of the  
 digital array or bitmap is written to the output memory at a  
 write speed, and the digital content from the output memory  
 is read at a read speed, said read speed being generally equal  
 to one half said write speed.

20. The method of claim 19 wherein the number of the  
 frames is four, and the number of the plurality of segments  
 is eight.

21. The method of claim 20 wherein the digital array or  
 bitmap corresponding to the two-dimensional electronic  
 image is composed of X lines, and wherein each one of the  
 plurality of segments corresponds to X/8 lines.

22. The method of claim 20 wherein the digital array or  
 bitmap corresponding to the two-dimensional electronic

45

image is composed of X lines, each of said lines having a first half and a second half, and wherein each one of the plurality of segments corresponds to X/8 of said first halves of said lines or X/8 of said second halves of said lines.

23. The method of claim 22 wherein the plurality of segments includes a first segment, a second segment adjacent to said first segment, a third segment adjacent to said first segment, a fourth segment adjacent to said third segment and said second segment, a fifth segment adjacent to said third segment, a sixth segment adjacent to said fifth segment and said fourth segment, a seventh segment adjacent to said fifth segment, and an eighth segment adjacent to said seventh segment and said sixth segment, each of said first segment, said third segment, said fifth segment, and said seventh segment being composed of X/8 of the first halves, each of said second segment, said fourth segment, said sixth segment, and said eighth segment being composed of X/8 of the second halves, wherein the number of the frames includes a first frame, a second frame, a third frame, and a fourth frame, said first segment and said second segment being stored in said first frame, said third segment and said fourth segment being stored in said second frame, said fifth segment and said sixth segment being stored in said third frame, and said seventh segment and said eighth segment being stored in said fourth frame.

24. The method of claim 22 wherein the plurality of segments includes a first segment, a second segment adjacent to said first segment, a third segment adjacent to said second segment, a fourth segment adjacent to said third segment, a fifth segment adjacent to said first segment, a sixth segment adjacent to said fifth segment and said second segment, a seventh segment adjacent to said sixth segment and said third segment, and an eighth segment adjacent to said seventh segment and said fourth segment, each of said first segment, said second segment, said third segment, and said fourth segment being composed of X/8 of the first halves, each of said fifth segment, said sixth segment, said seventh segment, and said eighth segment being composed of X/8 of the second halves, wherein the number of the frames includes a first frame, a second frame, a third frame, and a fourth frame, said first segment and said second segment being stored in said first frame, said third segment and said fourth segment being stored in said second frame, said fifth segment and said sixth segment being stored in said third frame, and said seventh segment and said eighth segment being stored in said fourth frame.

25. The method of claim 24 wherein the first frame and the second frame are located in sequence and adjoining one another on the storage medium.

26. The method of claim 24 wherein the first frame and the third frame are located in sequence and adjoining one another on the storage medium.

27. The method of claim 14 wherein the method further comprises the step of: seaming the plurality of segments together prior to or during the step of writing the corresponding digital content to the input memory such that the digital array or bitmap corresponding to the two-dimensional electronic image is reconstituted in the input memory.

28. The method of claim 27 wherein the digital array or bitmap corresponding to the two-dimensional electronic image is composed of a plurality of lines, an individual line of said plurality of lines including a first half and a second half, and wherein seaming the plurality of segments together comprises:

reading either the first half of said individual line plus a leading portion of said second half of said individual line or said second half of said individual line plus a trailing portion of said first half of said individual line;

46

comparing said leading portion of said second half of said individual line with said trailing portion of said first half of said individual line to determine an offset between said first half and said second half of said individual line; and

adjusting said first half or said second half to eliminate said offset.

29. The method of claim 28 wherein the leading portion and the trailing portion of the individual line are disposed in two different ones of the plurality of segments and in one of the frames.

30. The method of claim 28 wherein the leading portion and the trailing portion of the individual line are disposed in two different ones of the plurality of segments and two different ones of the frames.

31. A method for the storage and retrieval of a two-dimensional electronic image expressible as a digital array or bitmap for the purpose of interactive document processing, said method being utilized in association with a storage device having a storage medium, said method comprising the steps of:

writing at least a portion of the digital array or bitmap corresponding to the two-dimensional electronic image to an output memory;

reading said digital content from said output memory;

converting said digital content to a first analog signal as said digital content is read from said output memory;

adding a plurality of image control signals to said first analog signal to define a format, said format defining a frame;

selectively transmitting said first analog signal over the transmission pathway to the remote location or storing said first analog signal as at least one said frame on the storage device, the storage device being capable of receiving and storing said first analog signal on the storage medium along with a multiplicity of like frames, the storage device further being capable of randomly accessing and selectively retrieving and outputting a separate analog signal corresponding to said first analog signal stored as at least one said frame from among said multiplicity of like frames, said separate analog signal including said plurality of image control signals;

retrieving said separate analog signal including said plurality of image control signals from the storage device;

stripping said plurality of image control signals from said separate analog signal retrieved from the storage device;

converting said separate analog signal retrieved from the storage device to a digital signal;

writing said digital signal to an input memory such that the digital array or bitmap corresponding to the two-dimensional electronic image is reconstituted in said input memory,

whereby the electronic image may subsequently be utilized in one or more document processing operations including the transformation, presentation, representation, transmission, or storage and retrieval of the electronic image,

wherein the two-dimensional electronic image is captured from a tangible document and converted to the digital array or bitmap, said method further comprising the steps of:

capturing an initial image in a capture device, said initial image corresponding to a camera raster composed of X scanning lines having a predetermined scan rate, said

47

capture device producing an analog output signal corresponding to sequentially outputting said X scanning lines at one half said predetermined scan rate, said analog output signal including a camera raster synchronization;

stripping said camera raster synchronization from said analog output signal;

converting said analog output signal to a digital input signal; and

writing said digital input signal to a capture memory to form the digital array or bitmap corresponding to the two-dimensional electronic image, said digital array being composed of 2X lines.

32. The method of claim 31 wherein the analog output signal from the capture device has an initial bandwidth, and wherein the method further comprises the steps of:

converting the initial bandwidth of the analog output signal to a resultant bandwidth prior to storing or transmitting the first analog signal; and

converting said resultant bandwidth to the initial bandwidth subsequent to retrieving the separate analog signal or receiving the first analog signal.

33. The method of claim 32 wherein the initial bandwidth is on the order of 12 Mhz or greater and the resultant bandwidth is on the order of 6 Mhz.

34. The method of claim 31 wherein the capture device is a camera having a normal operating scan rate for producing the camera raster composed of the X scanning lines having the predetermined scan rate, and wherein the method further comprises the step of:

driving the camera at an actual scan rate generally equal to one half the normal operating scan rate.

35. The method of claim 31 wherein the method further comprises the step of:

multiplexing the digital input signal prior to or during the step of writing the digital input signal to the capture memory.

36. The method of claim 31 wherein the capture memory and the input memory are defined by one semiconductor memory.

37. The method of claim 31 wherein the initial image is composed of 1024 pixels in a vertical direction by 1000 pixels in a horizontal direction, and wherein the capture memory is a 1024 by 2048 bit memory.

38. The method of claim 31 wherein the capture memory is composed of two banks of eight 128 kbit memory.

39. An apparatus for the storage and retrieval of a two-dimensional electronic image existing as a digital array or bitmap in a memory for the purpose of interactive document processing by a user, said apparatus comprising:

a computer, said computer including a communications pathway, an electronic image processing interface operatively connected to said communications pathway, and the memory, the memory being operatively connected to said communications pathway;

a storage device, said storage device being operatively connected to said electronic image processing interface and including a storage medium capable of receiving and storing an analog signal in a designated form in a predetermined format defining a frame; and

said electronic image processing interface being capable of either:

selectively storing the two-dimensional electronic image to the storage device by converting the digital array or bitmap in the memory to a first analog signal

48

as said digital content is being read from the memory, and adding a plurality of image control signals to said first analog signal in the predetermined format corresponding to at least one of the frames, or

selectively retrieving the two-dimensional electronic image from the storage device by retrieving the at least one of the frames as a separate analog signal including said plurality of image control signals from said storage device, stripping said image control signals from said separate analog signal, converting said separate analog signal to a digital signal, and writing said digital signal to the memory such that the digital array or bitmap corresponding to the two-dimensional electronic image is reconstituted in the memory,

said apparatus further comprising:

a transmission network said transmission network being operatively connected to the electronic image processing interface, wherein the transmission network is capable of transmitting the first analog signal including the plurality of image control signals to a remote location whereby the first analog signal may be received by a second computer having a second electronic image processing interface and a second memory, said second electronic image processing interface being at least capable of stripping said plurality of image control signals from said first analog signal, converting said first analog signal to a digital signal, writing said digital signal to said second memory associated with said second electronic image processing interface such that the digital array or bitmap corresponding to the two-dimensional electronic image is reconstituted in the second memory of the second computer.

40. The apparatus of claim 39 wherein the electronic image processing interface is capable of both selectively storing the two dimensional electronic image to the storage device and selectively retrieving the two-dimensional electronic image from the storage device.

41. The apparatus of claim 39 wherein the storage device is operatively connected to the electronic image processing interface through the communications pathway.

42. The apparatus of claim 39 wherein the electronic image processing interface is further capable of sectioning the digital array or bitmap corresponding to the two-dimensional electronic image in the memory into a plurality of segments.

43. The apparatus of claim 42 wherein the transmission network includes a plurality of transmission pathways and each one of the plurality of segments may be transmitted in parallel over said plurality of transmission pathways.

44. The apparatus of claim 42 wherein the transmission network includes at least one transmission pathway having a transmission bandwidth, and wherein the number of the plurality of segments is determined such that the first analog signal corresponding to each of said plurality of segments is within said transmission bandwidth.

45. The apparatus of claim 39 wherein the first analog signal has an initial bandwidth, and wherein the electronic image processing interface is further capable of converting said initial bandwidth to a resultant bandwidth.

46. The apparatus of claim 45 wherein the transmission network includes at least one transmission pathway having a transmission bandwidth, and wherein the resultant bandwidth is within said transmission bandwidth.

47. The apparatus of claim 39 wherein the transmission network is operatively connected to the electronic image processing interface through the communications pathway.

## 49

48. The apparatus of claim 39 wherein the designated form in which the analog signal is stored is on a digital medium such as an optical laser disc.

49. A method for the storage and retrieval of a two-dimensional electronic image expressible as a digital array or bitmap for the purpose of interactive document processing, said method being utilized in association with a storage device having a storage medium, said method comprising the steps of:

writing at least a portion of the digital array or bitmap corresponding to the two-dimensional electronic image to an output memory;

reading said digital content from said output memory;

converting said digital content to a first analog signal as said digital content is read from said output memory;

adding a plurality of image control signals to said first analog signal to define a format, said format defining a frame;

selectively transmitting said first analog signal over the transmission pathway to the remote location or storing said first analog signal as at least one said frame on the storage device, the storage device being capable of receiving and storing said first analog signal on the storage medium along with a multiplicity of like frames, the storage device further being capable of randomly accessing and selectively retrieving and outputting a separate analog signal corresponding to said first analog signal stored as at least one said frame from among said multiplicity of like frames, said separate analog signal including said plurality of image control signals;

## 50

retrieving said separate analog signal including said plurality of image control signals from the storage device; stripping said plurality of image control signals from said separate analog signal retrieved from the storage device;

converting said separate analog signal retrieved from the storage device to a digital signal;

writing said digital signal to an input memory such that the digital array or bitmap corresponding to the two-dimensional electronic image is reconstituted in said input memory,

whereby the electronic image may subsequently be utilized in one or more document processing operations including the transformation, presentation, representation, transmission, or storage and retrieval of the electronic image,

wherein the first analog signal has an initial bandwidth, and wherein the method further comprises the steps of:

converting the initial bandwidth of the first analog signal to a resultant bandwidth prior to storing or transmitting the first analog signal; and

converting said resultant bandwidth to the initial bandwidth subsequent to retrieving the separate analog signal or receiving the first analog signal.

50. The method of claim 49 wherein the initial bandwidth is on the order of 12 Mhz or greater and the resultant bandwidth is on the order of 6 Mhz or less.

51. The method of claim 1 wherein both the output memory and the input memory are defined by one semiconductor memory.

\* \* \* \* \*



US005748805A

**United States Patent** [19]

Withgott et al.

[11] Patent Number: **5,748,805**[45] Date of Patent: **May 5, 1998**

[54] **METHOD AND APPARATUS FOR SUPPLEMENTING SIGNIFICANT PORTIONS OF A DOCUMENT SELECTED WITHOUT DOCUMENT IMAGE DECODING WITH RETRIEVED INFORMATION**

4,994,987 2/1991 Baldwin ..... 364/518  
 4,996,707 2/1991 O'Malley et al. .... 379/100  
 5,010,581 4/1991 Kanno ..... 382/56  
 5,048,099 9/1991 Lee ..... 382/282

(List continued on next page.)

[75] Inventors: **M. Margaret Withgott**, Los Altos, Calif.; **William Newman**, Cambridge, England; **Steven C. Bagley**, Palo Alto, Calif.; **Daniel P. Huttenlocher**, Ithaca, N.Y.; **Ronald M. Kaplan**, Palo Alto, Calif.; **Todd A. Cass**, Cambridge, Mass.; **Per-Kristian Halvorsen**, Los Altos, Calif.; **John Seely Brown**, San Francisco, Calif.; **Martin Kay**, Menlo Park, Calif.

**FOREIGN PATENT DOCUMENTS**

0 364 179 4/1990 European Pat. Off. .  
 0 364 180 4/1990 European Pat. Off. .  
 0 398 185 11/1990 European Pat. Off. .  
 62-154845 7/1987 Japan .  
 4-77965 3/1992 Japan .

**OTHER PUBLICATIONS**

Dan S. Bloomberg, "Multiresolution Morphological Approach to Document Image Analysis"; First International Conference on Document Analysis and Recognition ICDAR 91; 30 Sep. -2 Oct. 1991; St. Malo, France; pp. 963-971.  
 "Blind readers can use machine to recognize all fonts"; *COMPUTER*, vol. 12, No. 2, 1979, pp. 97-98.  
 "A Business Intelligence System" by H.P. Luhn, IBM Journal, Oct. 1958.  
 "Introduction to Modern Information Retrieval" by Salton and McGill, Chapter 2, pp. 24-51, McGraw-Hill, Inc. (1983).

*Primary Examiner*—Yon J. Couso  
*Attorney, Agent, or Firm*—Oliff & Berridge, PLC

[73] Assignee: **Xerox Corporation**, Stamford, Conn.[21] Appl. No.: **272,452**[22] Filed: **Jul. 11, 1994****Related U.S. Application Data**

[63] Continuation of Ser. No. 795,419, Nov. 19, 1991, abandoned.

[51] Int. Cl.<sup>6</sup> ..... **G06K 9/03**[52] U.S. Cl. .... **382/306; 707/512**

[58] Field of Search ..... **382/173, 308,**  
**382/305, 317, 312, 162-165, 309-311,**  
**306, 180, 175; 364/518, 419; 395/104,**  
**146, 147; 707/512**

**References Cited****U.S. PATENT DOCUMENTS**

4,475,239 10/1984 Van Raamsponk ..... 382/309  
 4,589,144 5/1986 Namba ..... 382/175  
 4,594,732 6/1986 Tsuji ..... 382/9  
 4,748,678 5/1988 Takeda et al. .... 382/306  
 4,847,912 7/1989 Tanaka et al. .... 382/9  
 4,965,763 10/1990 Zamora ..... 364/900  
 4,972,349 11/1990 Kleinberger ..... 364/900  
 4,985,863 1/1991 Fujisawa et al. .... 364/900  
 4,985,930 1/1991 Takeda et al. .... 382/56

**[57] ABSTRACT**

A method and apparatus for applying morphological image criteria that identify image units in an undecoded document image having significant information content, and for retrieving related data that supplements the document either from elsewhere within the document or a source external to the document. The retrieved data can result from character code recognition or template matching of the identified significant image units, or the retrieved data can result directly from an analysis of the morphological image characteristics of the identified significant image units. A reading machine can allow a user to browse and select documents or segments thereof, and to obtain interactive retrieval of documents and supplemental data.

**31 Claims, 5 Drawing Sheets****1.3.1.1 Les différentes catégories d'éditeurs interactifs**

Les éditeurs de texte interactifs peuvent être classés en quatre catégories: les éditeurs orientés par le système, les éditeurs orientés par la structure, les éditeurs de texte ou processeurs de texte, et les éditeurs/formateurs à composition graphique.

Les premiers correspondent aux outils qui se développent le plus souvent autour de la programmation. Ils sont en général liés à un langage de programmation et à ses caractéristiques. Ils permettent au programmeur d'augmenter sa productivité en l'aidant à présenter ses programmes correctement et en opérant des corrections syntaxiques qui réduisent le temps de mise au point. Au-delà de leurs fonctionnalités de traitement de texte, ils intègrent en fait certaines fonctions d'usage habituelles avec un compilateur ou un interpréteur.

Les seconds cherchent à exploiter la structure naturelle des documents. Ainsi une lettre est en soi parties modélisable sous la forme d'un arbre contenant: l'en-tête, le date, le nom et l'adresse du destinataire, le corps de la lettre, des paragraphes, des sous-paragraphes, une formule de politesse, une signature,

can  
composants  
texte  
structure  
graph  
recherche  
beyond  
distributed  
computer  
work  
Thun  
tree  
handling  
destination  
body  
patterns

5,748,805

Page 2

---

U.S. PATENT DOCUMENTS

5,048,109	9/1991	Bloomberg et al. ....	382/50	5,077,668	12/1991	Doi .....	364/419
5,050,218	9/1991	Ikeda et al. ....	382/1	5,175,684	12/1992	Chong .....	364/419
5,058,189	10/1991	Kanno .....	382/282	5,195,032	3/1993	Mastui et al. ....	364/419
				5,222,160	6/1993	Sakai et al. ....	382/57
				5,495,581	2/1996	Tsai .....	395/154

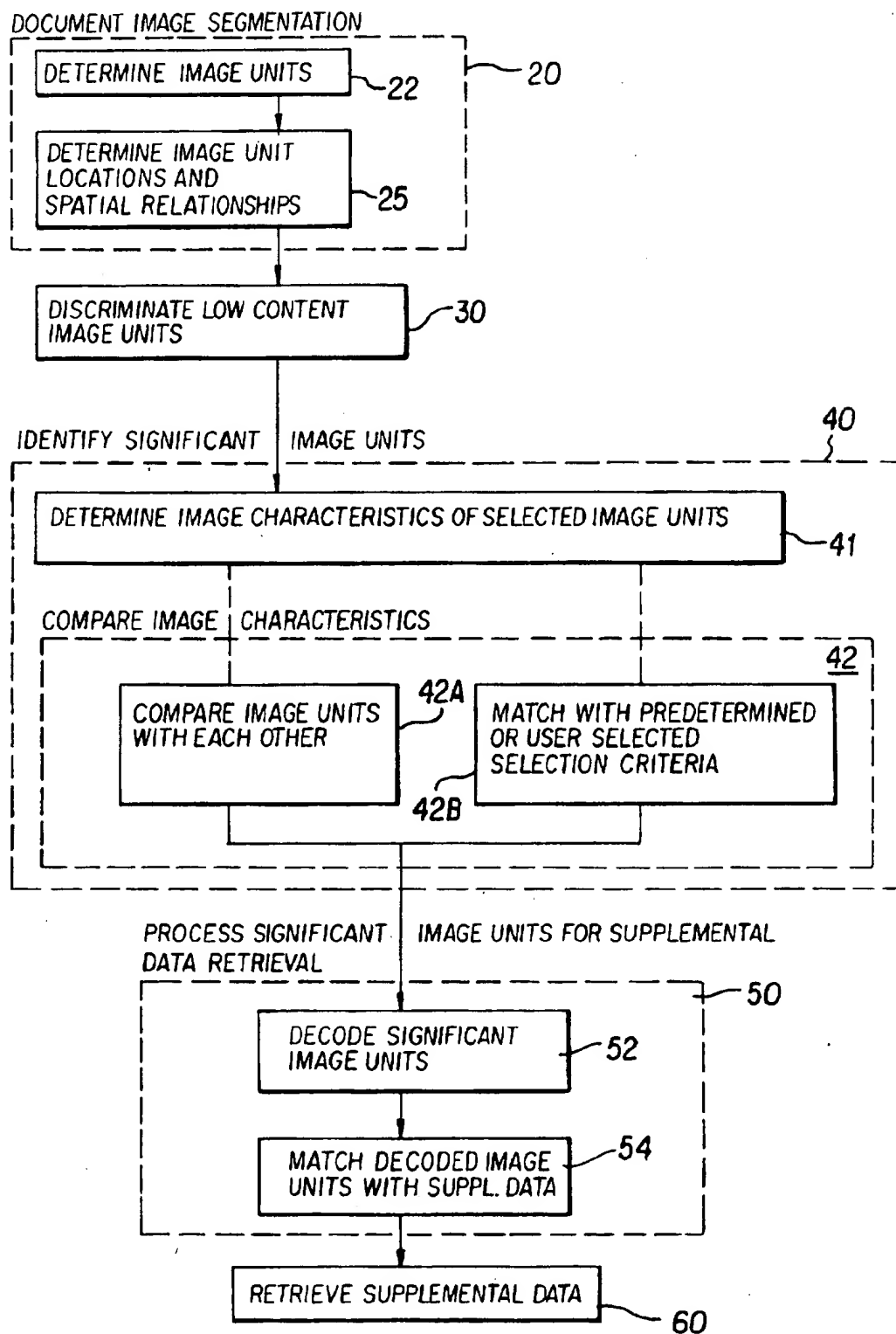


FIG. 1A

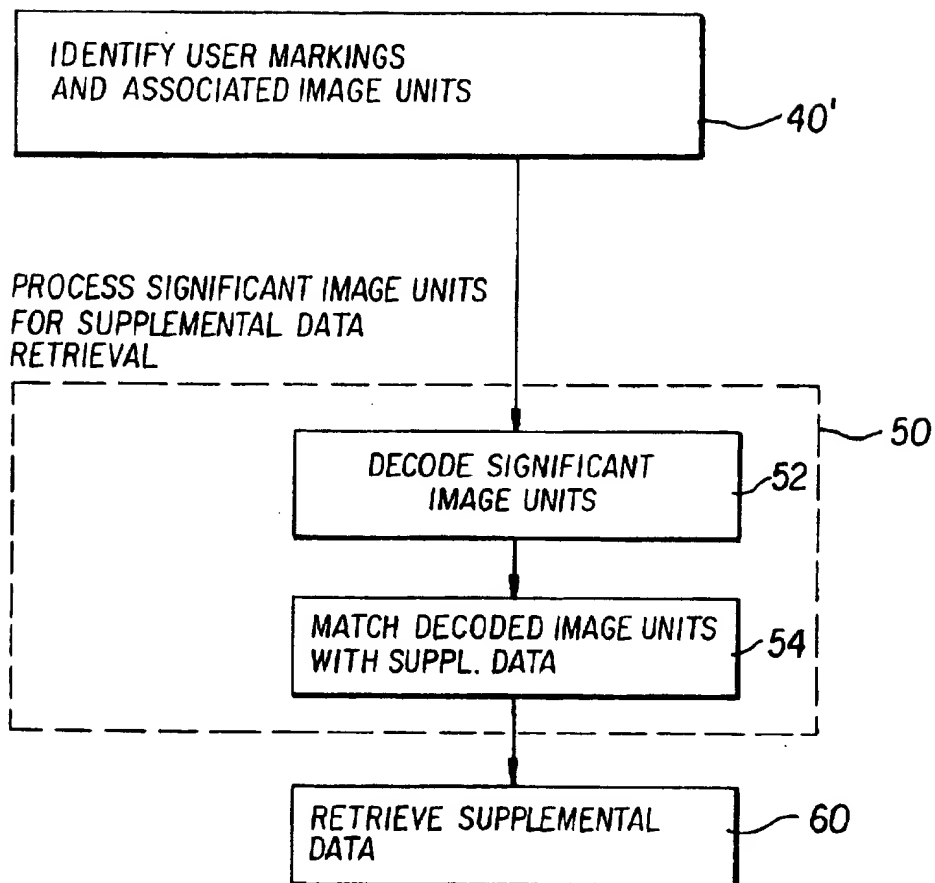


FIG. 1B

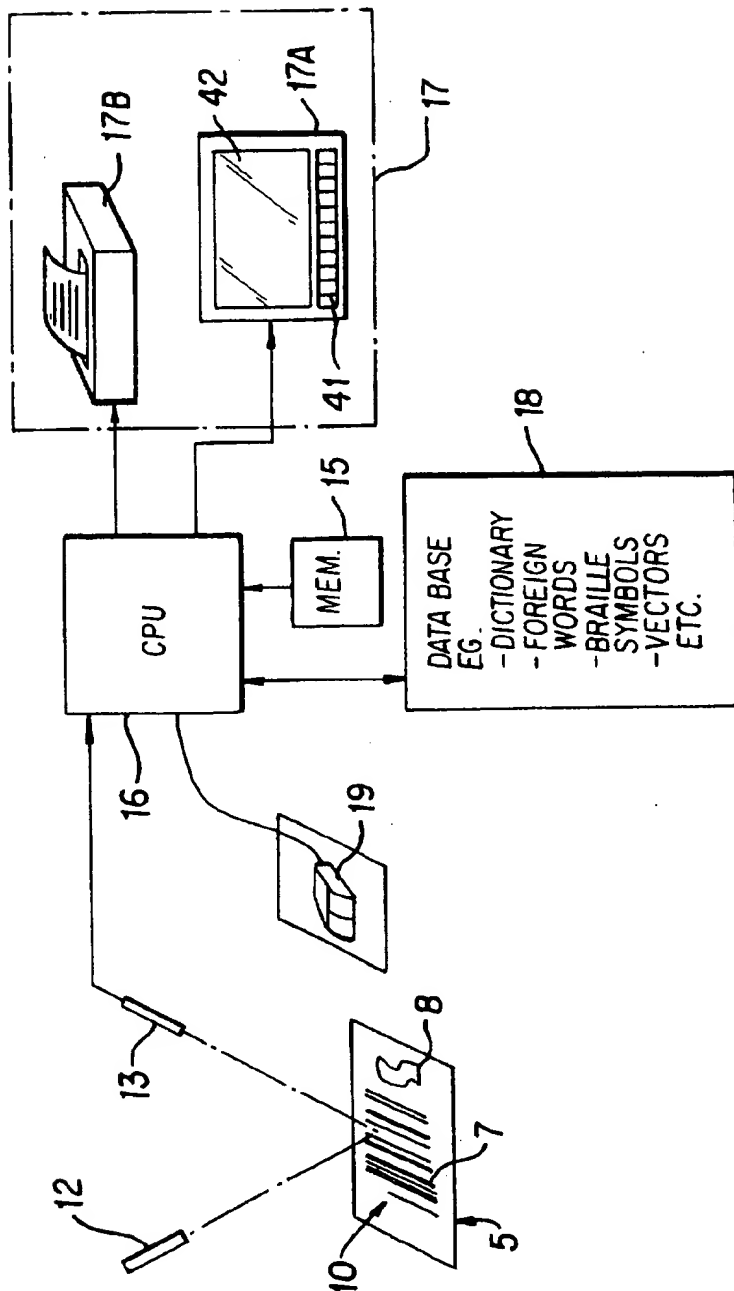


FIG. 2

### 1.3.1.1 Les différentes catégories d'éditeurs interactifs

Les éditeurs de texte interactifs peuvent être classés en quatre catégories: les éditeurs orientés par la syntaxe, les éditeurs orientés par la structure, les éditeurs de texte ou processeurs de texte, et les éditeurs/formateurs à composants graphiques.

Les premiers correspondant aux outils qui se développent le plus souvent autour de la programmation. Ils sont en general liés à un langage de programmation et à ses caractéristiques. Ils permettent au programmeur d'accroître sa productivité en l'aidant à présenter ses programmes correctement et en opérant des contrôles syntaxiques qui réduisent le temps de mise au point. Au delà de leurs fonctionnalités de traitement de texte, ils intègrent en fait certaines fonctions dévolues habituellement au compilateur ou à l'interpréteur.

Les seconds cherchent à exploiter la structure naturelle des documents. Ainsi une lettre est en sous parties modélisable sous la forme d'un arbre contenant: l'en-tête, la date, le nom et l'adresse du destinataire, le corps de la lettre, des paragraphes, des sous-paragraphes, une formule de politesse, une signature,

FIG. 3

### 1.3.1.1 Les différentes catégories d'éditeurs interactifs

Les éditeurs de texte interactifs peuvent être classés en quatre catégories: les éditeurs orientés par la syntaxe, les éditeurs orientés par la structure, les éditeurs de texte ou processeurs de texte, et les éditeurs/formateurs à composants graphiques.

Les premiers correspondant aux outils qui se développent le plus souvent autour de la programmation. Ils sont en général liés à un langage de programmation et à ses caractéristiques. Ils permettent au programmeur d'accroître sa productivité en l'aidant à présenter ses programmes correctement et en opérant des contrôles syntaxiques qui réduisent le temps de mise au point. Au dela de leurs fonctionnalités de traitement de texte, ils intègrent en fait certaines fonctions dévolues habituellement au compilateur ou à l'interpréteur.

Les seconds cherchent à exploiter la structure naturelle des documents. Ainsi une lettre est en sous parties modélisable sous la forme d'un arbre contenant: l'en-tête, la date, le nom et l'adresse du destinataire, le corps de la lettre, des paragraphes, des sous-paragraphes, une formule de politesse, une signature,

45  
can  
components  
tools  
often connected  
grow  
reduce put  
beyond  
delegated  
compiler  
seek  
Thus under/sub-  
tree heading  
destination body  
politeness

FIG. 4

# METHOD AND APPARATUS FOR SUPPLEMENTING SIGNIFICANT PORTIONS OF A DOCUMENT SELECTED WITHOUT DOCUMENT IMAGE DECODING WITH RETRIEVED INFORMATION

This is a Continuation of application Ser. No. 07/795,419 filed Nov. 19, 1991, now abandoned.

## BACKGROUND OF THE INVENTION

### 1. Cross-References to Related Applications

The following concurrently filed and related U.S. Patent Applications are hereby cross referenced and incorporated by reference in their entirety. "Method and Apparatus for Determining Boundaries of Words in Text" to Huttenlocher et al., U.S. patent application Ser. No. 07/794,392 now U.S. Pat. No. 5,321,770.

"Detecting Function Words Without Converting A Document to Character Codes" to Bloomberg et al., U.S. patent application Ser. No. 07/794,190 now abandoned.

"A Method of Deriving Wordshapes for Subsequent Comparison" to Huttenlocher et al., U.S. patent application Ser. No. 07/794,391 now abandoned.

"Method and Apparatus for Determining the Frequency of Words in a Document without Document Image Decoding" to Cass et al., U.S. patent application Ser. No. 07/796,173 now U.S. Pat. No. 5,208,969.

"Optical Word Recognition By Examination of Word Shape" to Huttenlocher et al., U.S. patent application Ser. No. 07/796,119 now abandoned.

"Method for Comparing Word Shapes" to Huttenlocher et al., U.S. patent application Ser. No. 07/795,169 now abandoned.

"A Method and Apparatus for Image Hand Markup Detection", U.S. patent application Ser. No. 07/794,275 now U.S. Pat. No. 5,201,011.

### 2. Field of the Invention

This invention relates to improvements in methods and apparatus for electronic document processing, and more particularly to improvements in methods and apparatus for automatically selecting semantically significant words, characters, images, or image segments in a document image without first decoding the document or otherwise understanding the information in the document, and augmenting the document with additional retrieved information relating to the selected words, characters, images, or image segments.

### 3. References and Background

One objective of computer based electronic document processing is to facilitate the user's access to and understanding of the information contained in a document or corpus of documents. However, in many cases, such as with a document in a language or form (e.g., non-Braille text for a blind user) which is foreign to the user, the user needs additional information or translation of the document in order to obtain any understanding of the document. In other cases, even after a user reaches a level of understanding about a document or group of documents, the user often desires to obtain supplemental information with which to enhance the user's understanding. However, locating semantically significant portions of a document, or among a collection of documents, for example, and evaluating the relative significance of such portions can be a very arduous and time-consuming task. The problem of selecting the most significant portions of documents and retrieving supplemen-

tal information related thereto is particularly difficult when dealing directly with bit mapped document images rather than with character code representations (such as ASCII for text images). In the past, perfectly recognizable scanned text has been treated as being interchangeable with electronically stored character code files, rather than as a special problem domain. However, in contrast to ASCII text files, which permit users to perform operations such as Boolean algebraic key word searches to locate text of interest, text information that is scanned without decoding is difficult to retrieve, without exhaustive viewing of each document, or without hand-crafting summaries of the documents for search purposes. Of course, document viewing or creation of document summaries requires extensive human effort.

Examples of retrieval techniques that rely upon locating useful terms in a document can be found in Salton and McGill, *Introduction To Modern Information Retrieval*, McGraw-Hill, Inc., 1983. Thus, techniques exist for computing key word matches, locating the most frequent noun phrases in a text, composing stop-lists of words which are not likely to be of interest to a user of an information retrieval system, and so on. Such techniques generally assume noise-free (perfectly recognizable) text.

### 4. References

U.S. Pat. No. 4,972,349 to Kleinberger describes a computerized information retrieval system and method formed of a textbase of texts of variable length and content. The texts are selected from the textbase on the basis of Boolean logic searches among key words associated with the texts. When a group is retrieved from such a search, the system automatically segregates the texts based on the presence or absence of a criteria-key keyword selected to segregate the texts into sub-groups. The same criteria key analysis can then be applied recursively to the subgroups. The resulting subgroups are then displayed to the user in a hierarchical display to illustrate the relationships among the texts. A string comparison routine is also described to search for similar keywords.

U.S. Pat. No. 4,985,863 to Fujisawa et al. describes a document storage and retrieval system and a method of document retrieval that stores a portion of characters for outputting and also stores the document in the form of an image for retrieving. A retrieval request for a text is made using a proper number of the text or a special symbol. The document image can then be retrieved and stored or decoded to character codes for outputting. Character recognition is performed to recognize a retrieval key of a document before retrieval, although the actual retrieval or transfer of the document does not require complete character recognition.

U.S. Pat. No. 5,010,581 to Kanno describes a data processing apparatus for retrieving abstract and original image data. The abstract image is an abbreviated form of the original image. The apparatus includes input means for inputting an original image; first memory means for temporarily storing the original image input by the input means; drafting means for drafting an abstract image of the original image; second memory means for storing the original and abstract images; and retrieval means for retrieving the abstract image based on retrieval data corresponding to both the original and abstract image. The second memory means stores the abstract image as a first page of the original image.

U.S. Pat. No. 4,996,707 to O'Malley et al. describes a computer system that includes a capability to receive and store graphic images from remote facsimile machines. The system includes software that can convert graphic images of textual material into an ASCII coded file so that either

keywords or the converted text may be converted to speech, giving the addressee-user the ability to review incoming facsimiles from a remote telephone. The system includes a relay capability, the capability to print on command and to originate facsimiles either from text files or scanned papers.

### SUMMARY OF THE INVENTION

Accordingly, it is an object of the invention to provide an improved method and apparatus for electronic document processing wherein supplemental data is retrieved for association with the electronic document which is relevant to significant portions of the document selected without decoding of the document.

It is another object of the invention to provide a method and apparatus of the type described that may perform retrieval operations based on morphological (structural) image characteristics of the document image to select the portions of the document on which the supplemental data retrieval is based, either with or without the supporting use of optical character recognition techniques to retrieve the supplemental data.

It is another object of the invention to provide a method and apparatus of the type described that may be used to provide supplemental information relating to a source document to be used in reading machines for the blind.

It is another object of the invention to provide a method and apparatus of the type described that may be used to provide translations for selected words in a source document.

In accordance with one aspect of the invention, a method for processing an undecoded document image is presented. According to the method, the document image is segmented into image units having information content without decoding of the document image. The significant image units on which the document supplementation is to be based are then identified, based solely on an evaluation of at least one morphological (structural) image characteristic of selected image units, or of hand-drawn document markings associated therewith. Supplemental data related to the identified significant image units are then retrieved, either with or without decoding of the identified significant image units in dependence on the form of the supplemental data.

The morphological image characteristics used to identify significant image units include image unit shape dimensions, typeface, font, location in the document image and frequency of image unit occurrence. In one embodiment, the significant image units are identified according to hand-drawn graphical markings placed on the document by the user adjacent word units of interest to the user, such as encircling or underscoring.

The retrieval method may be used to retrieve, for example, foreign language translation data corresponding to the selected image units, or Braille versions of the selected image units for print out. The supplemental data may also take the form of a different mode of display, such as speech synthesized verbal output of the selected image units.

In accordance with another aspect of the invention, an apparatus for retrieving data to supplement a document is presented. The apparatus includes a scanner for scanning the document image, and means for segmenting the document image into image units. Means are provided for classifying selected image units as significant image units, and means are provided for retrieving supplemental data related to the significant image units for utilization with the document. The apparatus may include a programmed digital computer to provide the means for segmenting the document image,

means for classifying significant image units, and the means for retrieving supplemental data.

These and other objects, features and advantages of the invention will be apparent to those skilled in the art from the following detailed description, when read in conjunction with the accompanying drawings and appended claims.

### BRIEF DESCRIPTION OF THE DRAWINGS

A preferred embodiment of the invention is illustrated in the accompanying drawings, in which:

FIG. 1A is a flow chart of a first embodiment of a method of the invention.

FIG. 1B is a flow chart of a second embodiment of a method of the invention.

FIG. 2 is a block diagram of an apparatus according to the invention for carrying out the method embodiments of either FIG. 1A or 1B.

FIG. 3 is an example of a text document on which selected terms have been underlined for identification and association with supplemental information, in accordance with the method embodiment of FIG. 1B.

FIG. 4 is an example of the document of FIG. 3 on which supplemental information has been associated in accordance with the method of the invention.

### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

In contrast to prior techniques, such as those described above, the invention is based upon the recognition that scanned image files and character code files exhibit important differences for image processing, especially in data retrieval. The method of a preferred embodiment of the invention capitalizes on the visual properties of text contained in paper documents, such as the presence or frequency of linguistic terms (such as words of importance like "important", "significant", "crucial", or the like) used by the author of the text to draw attention to a particular phrase or a region of the text; font and type face variations used to emphasize significant words, formatting conventions, and so on. Two preferred embodiments of the method of the invention are illustrated in the flow charts of FIGS. 1A and 1B, and an apparatus for performing both methods is shown in FIG. 2. For the sake of clarity, the invention will be described with reference to the processing of a single document. However, it will be appreciated that the invention is applicable to the processing of a corpus of documents containing a plurality of documents.

The invention provides a method and apparatus for retrieving data to supplement the data or text in, for example, a scanned document. However, the invention is not limited to systems utilizing document scanning. Rather, other systems such as a bitmap workstation (i.e., a workstation with a bitmap display) or a system using both bitmapping and scanning would work equally well for the implementation of the methods and apparatus described herein.

The supplementing data may be retrieved from a wide number of sources, for example, from internal data in another portion or page of the source document, or from external data, such as an on-line data base. The retrieved data can be presented in any number of ways. For example, it can be presented as marginal notes on the source document, i.e., the document containing material to be supplemented, adjacent the source document portion(s) to which it pertains; printed as footnotes on the source document; printed on separate documents or in special formats

(e.g., Braille); presented in a different form of display, such as a video display or synthesized voice output; or in some other appropriate manner.

Still more particularly, in accordance with one aspect of the invention (FIG. 1B), selected portions of the source document can be denoted by a user by hand-marking the source document. Due to the varied applications for the invention, the marking may take various forms. For example, a user may underline or encircle or otherwise highlight selected words to be supplemented. An example of such selection is illustrated in FIG. 3, in which selected words in a text 7 have been marked by underlining 11.

The selection, on the other hand, may take a more automatic form (FIG. 1A), for example, in which significant portions of the source document are targeted and automatically identified by a general purpose digital computer or the like according to one or more morphological image characteristics which are predetermined or selected by the user.

With reference first to FIG. 2, the method is performed on an electronic image of an original document 5 (e.g., a bitmap image or a scanned image), which may include lines of text 7, titles, drawings, FIG. 8, or the like, contained in one or more sheets or pages of paper 10 or other tangible form. The electronic document image to be processed is created in any conventional manner, for example, by an input means, such as an optical scanner 12 and sensor 13 as shown, a document copier machine scanner, a Braille reading machine scanner, a bitmap workstation, an electronic beam scanner or the like. Such means are well known in the art, and thus are not described in detail herein. An output derived from, for example, a scanner sensor 13 is digitized to produce undecoded bit mapped image data representing the document image for each page of the document, which data is stored, for example, in a memory 15 of a special or general purpose digital computer 16. The output from the computer 16 is delivered to an output device 17, such as, for example, a memory or other form of storage unit; an output display 17A as shown, which may be, for instance, a CRT display; a printer device 17B as shown, which may be incorporated in a document copier machine or a Braille or standard form printer; a facsimile machine, speech synthesizer or the like.

The more automatic embodiment of the method of the invention, in which the image characteristics of selected image units are evaluated, will now be described with reference to FIG. 1A. The first phase of the image processing technique of the invention involves a low level document image analysis in which the document image for each page is segmented into undecoded information containing image units (step 20) using conventional image analysis techniques; or, in the case of text documents, preferably using the bounding box method described in concurrently filed U.S. patent application Ser. No. 07/794,392 filed concurrently herewith by Huttenlocher and Hopcroft, and entitled "Method and Apparatus for Determining Boundaries of Words in Text".

Another method for finding word boxes is to close the image with a horizontal SE that joins characters but not words, followed by an operation that labels the bounding boxes of the connected image components (which in this case are words). The process can be greatly accelerated by using 1 or more threshold reductions (with threshold value 1), that have the effect both of reducing the image and of closing the spacing between the characters. The threshold reduction(s) are typically followed by a closing with a small horizontal SE. The connected component labeling operation is also done at the reduced scale, and the results are scaled

up to full size. The disadvantage of operating at reduced scale is that the word bounding boxes are only approximate; however, for many applications the accuracy is sufficient. The described method works fairly well for arbitrary text fonts, but in extreme cases, such as large fixed width fonts that have large inter-character separation or small variable width fonts that have small inter-word separation, mistakes can occur. The most robust method chooses a SE for closing based on a measurement of specific image characteristics. This requires adding the following two steps:

(1) Order the image components in the original or reduced (but not closed) image in line order, left to right and top to bottom.

(2) Build a histogram of the horizontal inter-component spacing. This histogram should naturally divide into the small inter-character spacing and the larger inter-word spacings. Then use the valley between these peaks to determine the size of SE to use for closing the image to merge characters but not join words.

Once the bounding boxes or word boxes are determined, the locations of and spatial relationships between the image units on a page can be determined (step 25). For example, an English language document image can be segmented into word image units based on the relative difference in spacing between characters within a word and the spacing between words. Sentence and paragraph boundaries can be similarly ascertained. Additional region segmentation image analysis can be performed to generate a physical document structure description that divides page images into labelled regions corresponding to auxiliary document elements like figures, tables, footnotes and the like. Figure regions can be distinguished from text regions based on the relative lack of image units arranged in a line within the region, for example. Using this segmentation, knowledge of how the documents being processed are arranged (e.g., left-to-right, top-to-bottom), and, optionally, other inputted information such as document style, a "reading order" sequence for word images can also be generated. The term "image unit" is thus used herein to denote an identifiable segment of an image such as a number, character, glyph, symbol, word, phrase or other unit that can be reliably extracted. Advantageously, for purposes of document review and evaluation, the document image is segmented into sets of signs, symbols or other elements, such as words, which together form a single unit of understanding. Such single units of understanding are characterized in an image as being separated by a spacing greater than that which separates the elements forming a unit.

Advantageously, a discrimination step 30 is next performed to identify the image units which have insufficient information content to be useful in evaluating the subject matter content of the document being processed. One preferred method for use with text documents is to use the morphological function or stop word detection techniques disclosed in the copending U.S. patent application Ser. No. 07/794,190 filed concurrently herewith by Bloomberg et al., and entitled "Detecting Function Words Without Converting A Document to Character Codes".

Next, in step 40, selected image units, e.g., the image units not discriminated in step 30, are evaluated, without decoding the selected image units or reference to decoded image data, based on an evaluation of predetermined or selected image characteristics of the image units. The evaluation entails a determination (step 41) of the morphological image characteristics and a comparison (step 42) of the determined image characteristics for each image unit either with the determined morphological image characteristics of the other

image units or with predetermined morphological image characteristics or morphological image characteristics selected by the user.

One preferred method for defining the image unit image characteristics to be evaluated is to use the word shape derivation techniques disclosed in the copending U.S. patent application Ser. No. 07/794,391 filed concurrently herewith by D. Huttenlocher and M. Hopcroft, and entitled "A Method for Deriving Wordshapes for Subsequent Comparison". As described in the aforesaid application, at least one, one-dimensional signal characterizing the shape of the word unit is derived; or an image function is derived defining a boundary enclosing the word unit, and the image function is augmented so that an edge function representing edges of the character string detected within the boundary is defined over its entire domain by a single independent variable within the closed boundary, without individually detecting and/or identifying the character or characters making up the word unit.

The determined morphological characteristic(s), e.g., the derived image unit shape representations, of each selected image unit are compared, as noted above (step 42), either with the determined morphological image characteristic(s) or derived image unit shape representations of the other selected image units (step 42A), or with predetermined/user-selected image characteristics to locate specific types of image units (step 42B). The determined morphological image characteristics of the selected image units are advantageously compared with each other for the purpose of identifying equivalence classes of image units such that each equivalence class contains most or all of the instances of a given image unit in the document, and the relative frequencies with which image units occur in a document can be determined, as is set forth more fully in the copending U.S. patent application Ser. No. 07/795,173 filed concurrently herewith by Cass et al. now abandoned, and entitled "Method and Apparatus for Determining the Frequency of Words in a Document with Document Image Decoding". Image units can then be classified or identified as significant according to the frequency of their occurrence, as well as other characteristics of the image units, such as their length. For example, it has been recognized that for business communications in English, a useful combination of selection criteria is to select the medium frequency word units.

It will be appreciated that the specification of the image characteristics for titles, headings, captions, linguistic criteria or other significance indicating visual features of a document image can be predetermined and selected by the user to determine the selection criteria defining a "significant" image unit. Comparing the image characteristics of the selected image units of the document image for matches with the image characteristics associated with the selection criteria permits the significant image units to be readily identified without any document decoding.

Any of a number of different methods of comparison can be used. One technique that can be used, for example, is by correlating the raster images of the extracted image units using decision networks, such technique being described for characters in a Research Report entitled "Unsupervised Construction of Decision Networks for Pattern Classification" by Casey et al., IBM Research Report, 1984, said Research Report being incorporated by reference herein.

Preferred techniques that can be used to identify equivalence classes of word units are the word shape comparison techniques disclosed in U.S. patent application Ser. Nos. 07/796,119 and 07/795,169 now abandoned, filed concur-

rently herewith by Huttenlocher and Hopcroft, and by Huttenlocher, Hopcroft and Wayner, respectively, and entitled, respectively, "Optical Word Recognition By Examination of Word Shape," and "Method for Comparing Word Shapes".

Depending on the particular application, and the relative importance of processing speed versus accuracy, for example, evaluations of different degrees of precision can be performed. For example, useful evaluations can be based on length, width and/or other measurement dimensions of the image unit (or derived image unit shape representation, e.g., the largest figure in a document image); the location of the image unit in the document (including any selected figure or paragraph of a document image, e.g., headings, initial figures, one or more paragraphs or figures), font, typeface, cross-section (a cross-section being a sequence of pixels of similar state in an image unit); the number of ascenders; the number of descenders; the average pixel density; the length of a top line contour, including peaks and troughs; the length of a base contour, including peaks and troughs; the location of image units with respect to neighboring image units; vertical position; horizontal inter-image unit spacing; and combinations of such classifiers.

Referring to FIG. 1B, the embodiment of the method of the invention in which significant image units are selected based on user hand-drawn markings placed on the document does not require an initial document image segmentation step. Instead, the morphological method for identifying hand-drawn graphical markings disclosed in copending U.S. application Ser. No. 07/794,275, filed concurrently herewith, by Bloomberg, and entitled "A Method and Apparatus for Image Hand Markup Detection", now U.S. Pat. No. 5,201,011 is preferably utilized (step 40') to identify the regions of the document image containing the user markings. This method also permits the image units associated with the user markings to be identified.

In instances in which multiple page documents are processed, each page is processed and the data held in the memory 15 (see FIG. 1), as described above. The entirety of the data can then be processed.

Through use of equipment such as illustrated in FIG. 2, the identified word units 11 are morphologically detected; that is, significant morphological (structural) image characteristics inherent in the image form of the word units are detected. The non-content based image recognition aspect of the invention allows image processing of documents to provide integral information about the documents without first converting text in the document to character codes. Data retrieval can be then provided to automatically and directly access supplemental information associated with the detected word units.

A salient feature provided by the method of the invention is the initial processing and identification of significant word units being accomplished without an accompanying requirement that the content of the word units be decoded. More particularly, to this stage in the process, the actual content of the word units is not required to be specifically determined. Thus, for example, in such applications as copier machines or electronic printers that can print or reproduce images directly from one document to another without regard to ASCII or other encoding/decoding requirements, image units can be identified and processed using one or more morphological image characteristics or properties of the image units. The image units of unknown content can then be further optically or electronically processed. One of the advantages that results from the ability to perform such

image unit processing without having to decode the image unit contents at this stage of the process is that the overall speed of image handling and manipulation can be significantly increased.

The second phase of the document analysis according to both method embodiments of the invention involves further processing (step 50) of the identified image units in connection with the supplemental data retrieval. The further processing can be accomplished using a number of different techniques, depending upon the particular application. For example, word units 11 (FIG. 3) that have been identified from the scanned document may be decoded (step 52) by optical character recognition techniques, which techniques are well known in the art and thus are not described herein in detail. The decoded word units are then matched (step 54) with associated supplemental data in a conventional manner. For instance, in one embodiment, the supplemental data is contained in a data base 18 (see FIG. 2) that may contain specific data pertaining to the specific decoded word units. Data base 18 may be, for example, a dictionary containing definitions of the decoded words, translations of foreign words, or cross-references to related documents. Alternatively, the supplemental data may be vectors or keys to particular data such as synthesized speech data, memory locations, etc.

The supplemental data is then retrieved and outputted (step 60) to an appropriate output device. In the embodiment exemplified in FIGS. 3 and 4, for example, the supplemental data is translated words 45 corresponding to the words underlined by the user. In this case, the translated words are outputted by adding them to the document image in a conventional manner so that they appear in the margin adjacent the line containing the words underlined by the user when the document image is printed or displayed, as shown in FIG. 4.

Thus, employing the method and apparatus of the invention, a "translating copier" machine, for example, may be constructed to assist a user in understanding documents written in foreign languages. Depending on the degree to which the user is familiar with the foreign language, the user may either mark difficult or unknown words in a printed copy of the document or portion thereof for translation supplementation, or enter through an appropriate user interface a request that all significant words in the document or document portion be automatically selected in accordance with either predetermined or user-selected significance criteria. The translating photocopier then either scans the marked-up copy of the document and identifies the marked word units in accordance with the above described FIG. 1B method embodiment, or evaluates the image characteristics of selected word units in the scanned document image pursuant to the user's request to identify significant word units in accordance with the above described FIG. 1A embodiment. The copier then retrieves the relevant translation supplemental data as described above, and prints a fresh copy of the document or document portion with translations of the underlined words in the margins opposite the underlined words, as shown in FIG. 4.

Another application of the document supplementation techniques of the invention is in reading machines for the blind. One embodiment supports the designation by a user of key words, for example, on a key word list to designate likely points of interest in a document. Using the user designated key words, occurrences of the word can be found in the document of interest by OCR techniques or the like, and regions of text forward and behind the key word can be retrieved and processed using the techniques described

above. Or, significant key words can be automatically determined using the morphological recognition techniques described above. The words thus identified as significant words or word units can then be decoded using optical character recognition techniques, for example, for retrieval of supplemental data which permits, for example, Braille versions of the significant words to be printed using a plastic-based ink printer associated with the reading machine. Alternatively, speech synthesized output devices can be employed to produce a voice output representation of the significant words as the supplemental data.

Once a document has been supplemented, the user may wish to return to the original source to have printed or hear a full text rendition. This may be achieved in a number of ways. One method is for a synthesizer or printer to provide source information, for example, "on top of page 2 is an article entitled . . ." The user would then return to the point of interest.

Two classes of apparatus extend this capability through providing the possibility of user interaction while the supplemental data is being communicated. One type of apparatus is a simple index marker. This can be, for instance, a hand held device with a button that the user depresses whenever he or she hears a title of interest, or, for instance, an N-way motion detector in a mouse 19 (FIG. 2) for registering a greater variety of commands. The reading machine records such marks of interest and returns to the original article after the supplemental data is communicated.

Another type of apparatus makes use of the technology of touch-sensitive screens. Such an apparatus operates by requiring the user to lay down a Braille summarization sheet 41 on a horizontal display. The user then touches the region of interest on the screen 42 in order to trigger either a full printout or synthesized reading. The user would then indicate to the monitor when a new page was to be processed.

It will be appreciated that the method of the invention reduces the amount of material presented to the user for evaluation, and thus is capable of circumventing many problems inherent in the use of current reading technology for the blind and others, such as the problems associated with efficient browsing of a document corpus, using synthesized speech, and the problems created by the bulk and expense of producing Braille paper copies, and the time and effort required to read such copies.

The method of the invention has been described above to perform document retrieval using conventional character recognition techniques, such as OCR, in conjunction with morphological identification techniques. It will be appreciated, however, that direct retrieval using only image characteristic word unit recognition techniques may be performed in the case of supplemental data which is also stored as bit mapped image data compatible with the image data of the source document to be supplemented.

Although the invention has been described and illustrated with a certain degree of particularity, it is understood that the present disclosure has been made only by way of example, and that numerous changes in the combination and arrangement of parts can be resorted to by those skilled in the art without departing from the spirit and scope of the invention, as hereinafter claimed.

We claim:

1. A method for electronically processing an electronic document image, comprising:

segmenting the document image into word units without decoding the document image, each word unit corresponding to a word in said document image;

identifying significant ones of said word units in accordance with at least one selected morphological image characteristic of said word units without determining their content other than said at least one selected morphological image characteristic; and

retrieving supplemental data related to the identified significant word units.

2. The method of claim 1 wherein said step of identifying significant word units includes a step of classifying said word units according to frequency of occurrence.

3. The method of claim 1 wherein said step of identifying significant word units includes a step of classifying said word units according to location within the document image.

4. The method of claim 1 wherein said at least one selected morphological image characteristic includes image characteristics defining word units having predetermined linguistic criteria.

5. The method of claim 1 wherein said at least one selected morphological image characteristic includes at least one of a word unit shape dimension, font, typeface, number of ascender elements, number of descender elements, pixel density, pixel cross-sectional characteristic, the location of word units with respect to neighboring word units, vertical position, horizontal inter-image unit spacing, and contour characteristic of said word units.

6. The method of claim 1 wherein said step of identifying significant word units comprises identifying word units having an associated hand-drawn marking created by a user.

7. The method of claim 1 wherein said step of retrieving supplemental data comprises retrieving foreign language data corresponding to said identified significant word units.

8. The method of claim 1 wherein said step of retrieving supplemental data comprises retrieving data from within the document.

9. The method of claim 1 wherein said step of retrieving supplemental data comprises retrieving data external to the document.

10. The method of claim 1 further comprising modifying the document image with retrieved data to provide a document annotation.

11. The method of claim 10 wherein the document annotation is in the form of marginal notes which are located in a margin of said modified document image.

12. The method of claim 1 wherein said step of retrieving supplemental data retrieves Braille versions of the identified significant word units.

13. The method of claim 12 further comprising outputting said retrieved Braille versions of the identified significant word units in printed form.

14. The method of claim 1 wherein said step of retrieving supplemental data for the document comprises retrieving synthesized speech versions of the identified significant word units.

15. A method for translating a selected word in a text document, comprising:

marking the selected word in the document text with a hand-drawn graphical notation;

scanning the text document to produce an undecoded scanned document image;

segmenting the document image into image segments without decoding of the document image;

evaluating the morphological image characteristics of the scanned document image to identify said graphical notation;

identifying the image unit associated with said identified graphical notation;

retrieving translation data related to said identified image unit; and

incorporating said retrieved translation data in said scanned document image.

16. The method of claim 15 wherein said retrieving step includes decoding the identified image unit and matching the decoded image unit with decoded data entries in a dictionary database.

17. A method for electronically processing an undecoded document image containing word text, comprising:

segmenting the document image into word image units without decoding the document image, each word image unit corresponding to a word in said document image;

evaluating selected word image units according to at least one morphological image characteristic of the selected word image units without determining their content other than said at least one morphological image characteristic to identify significant word image units;

retrieving supplemental data related to the identified significant word image units; and

outputting said retrieved supplemental data.

18. An apparatus for retrieving data to supplement a document, comprising:

means for inputting the document to produce an undecoded document image;

means for segmenting the document image into word units having undecoded information content without decoding the document image, each word unit corresponding to a word in said document image;

means for evaluating selected word units according to at least one morphological image characteristic of said selected word units, without determining their content other than said at least one morphological image characteristic, to identify significant word units;

means for retrieving supplemental data related to the identified significant word units; and

an output device which utilizes the retrieved supplemental data.

19. The apparatus of claim 18 wherein said means for segmenting the document image, said means for identifying significant word units, and said means for retrieving supplemental data comprise a programmed digital computer.

20. The apparatus of claim 18, wherein the means for inputting the document comprises a bitmap workstation.

21. The apparatus of claim 20 wherein said retrieved supplemental data is translation data.

22. The method of claim 1, wherein said step of retrieving supplemental data comprises retrieving data from an online database.

23. The apparatus of claim 22 wherein said output device is a Braille printer for producing document copies in Braille format.

24. The apparatus of claim 22 wherein said output device is a speech synthesizer for producing synthesized speech output corresponding to said identified significant word units.

25. An apparatus for retrieving data to supplement a document, comprising:

a copier machine having a scanner for inputting the document to produce an undecoded document image;

means for segmenting the document image into word units having undecoded information content without decoding the document image, each word unit corresponding to a word in said document image;

13

means for evaluating selected word units according to at least one morphological image characteristic of said selected word units, without determining their content other than said at least one morphological image characteristic, to identify significant word units;

means for retrieving supplemental data related to the identified significant word units;

said copier machine including an output device for producing printed documents; and

means for incorporating said retrieved supplemental data into the document image for printing as part of a printed document by said output device.

26. An apparatus for retrieving data to supplement a document, comprising:

a Braille reading machine including a means for inputting the document to produce an undecoded document image;

means for segmenting the document image into word units having undecoded information content without decoding the document image, each word unit corresponding to a word in said document image;

means for evaluating selected word units according to at least one morphological image characteristic of said

14

selected word units, without determining their content other than said at least one morphological image characteristic, to identify significant word units;

means for retrieving supplemental data related to the identified significant word units; and

said Braille reading machine including an output device which utilizes the retrieved supplemental data to produce an output intelligible to a blind user.

27. The method of claim 1, further comprising printing said retrieved supplemental data with a copier machine.

28. The method of claim 27, wherein, prior to performing said segmenting step, said electronic document image is input to said copier machine.

29. The method of claim 28, wherein said electronic document image is input with a scanner of said copier machine.

30. The method of claim 15, wherein said text document is scanned using a scanner of a copier machine.

31. The method of claim 30, further comprising printing the scanned document image incorporating said retrieved translation data with a printing device of said copier machine.

\* \* \* \* \*



US005664030A

## United States Patent [19]

Iizuka

[11] Patent Number: 5,664,030

[45] Date of Patent: Sep. 2, 1997

102(b)

[54] METHOD AND APPARATUS FOR  
RECORDING/REPRODUCING MESH  
PATTERN DATA

[75] Inventor: Nobuo Iizuka, Hamura, Japan

[73] Assignee: Casio Computer Co., Ltd., Tokyo,  
Japan

[21] Appl. No.: 451,544

[22] Filed: May 26, 1995

## Related U.S. Application Data

[62] Division of Ser. No. 869,012, Apr. 14, 1992, Pat. No. 5,454,054, which is a division of Ser. No. 530,630, May 30, 1990, Pat. No. 5,153,928.

## [30] Foreign Application Priority Data

Jun. 9, 1989	JP	Japan	1-145255
Jun. 9, 1989	JP	Japan	1-145256
Jun. 9, 1989	JP	Japan	1-145257

[51] Int. Cl.<sup>6</sup> G06K 9/20

[52] U.S. Cl. 382/321

[58] Field of Search 382/112, 182,  
382/183, 184, 306, 309, 310, 311, 321;  
235/432, 436, 456, 437, 494, 462, 438

## [56] References Cited

## U.S. PATENT DOCUMENTS

3,780,271	12/1973	Sharkitt et al.	235/61.11 E
3,987,226	9/1976	Bunker et al.	340/146.1
4,228,716	10/1980	Linford	84/1.18
4,322,754	3/1982	Mason	358/296
4,422,361	12/1983	Ishii et al.	84/1.18
4,437,378	3/1984	Ishida et al.	84/1.18
4,464,966	8/1984	Ishida	84/1.03
4,533,825	8/1985	Yamada	235/463
4,641,357	2/1987	Satoh	382/61
4,728,783	3/1988	Brass et al.	235/456
4,783,838	11/1988	Matsunawa	382/51
4,788,598	11/1988	Ochi et al.	358/260
4,794,239	12/1988	Allais	235/462
4,924,078	5/1990	Sant'Anselmo et al.	235/494
4,934,846	6/1990	Gilham	400/104

4,939,354	7/1990	Priddy et al.	235/456
4,958,337	9/1990	Yamanaka et al.	369/58
4,972,475	11/1990	Sant'Anselmo	380/54

(List continued on next page.)

## FOREIGN PATENT DOCUMENTS

0384955	9/1990	European Pat. Off.
0439682	8/1991	European Pat. Off.
53-73026	6/1978	Japan

## OTHER PUBLICATIONS

I/O 1988, 5 pp. 121-125.

Jas Journal 1987, vol. 27, No. 10, pp. 39-52.

Nihon Keizai Shimbun newspaper article, published by Nihon Keizai Shimbun, Tokyo, Japan, Apr. 10, 1992.

Nihon Keizai Shimbun article; Apr. 10, 1992.

Article from the Denpa Shinbun newspaper, Tokyo, Japan, dated Jun. 8, 1992.

Itochu Electronics Co., Ltd., brochure, distributed at the exhibition, "Scan-Tech Japan 1992".

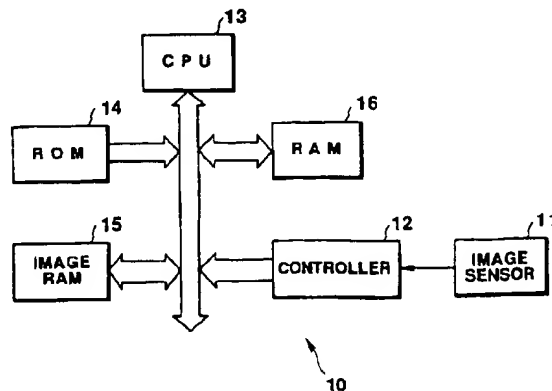
Primary Examiner—Yon J. Couso

Attorney, Agent, or Firm—Frishauf, Holtz, Goodman, Langer & Chick

## [57] ABSTRACT

Recorded data is read by an image sensor from a recording medium on which the data is two-dimensionally recorded in a mesh pattern, and the read data is stored in an image RAM. The stored data is subjected to data decoding and error correction to reproduce target data. A scanning reference pattern of a mesh pattern recorded on the recording medium serves as a guide when the stored recorded data read by the image sensor is reproduced by scanning in the image RAM. A reproduction apparatus has a capability of coping with a partial destruction of the scanning reference pattern. Error checking codes are added to the recording medium in addition to the target data. These items of data are recorded as an encoded image after being subjected to scrambling processing and randomization processing. The reproduction apparatus can easily cope with an error when the encoded image is reproduced.

2 Claims, 32 Drawing Sheets

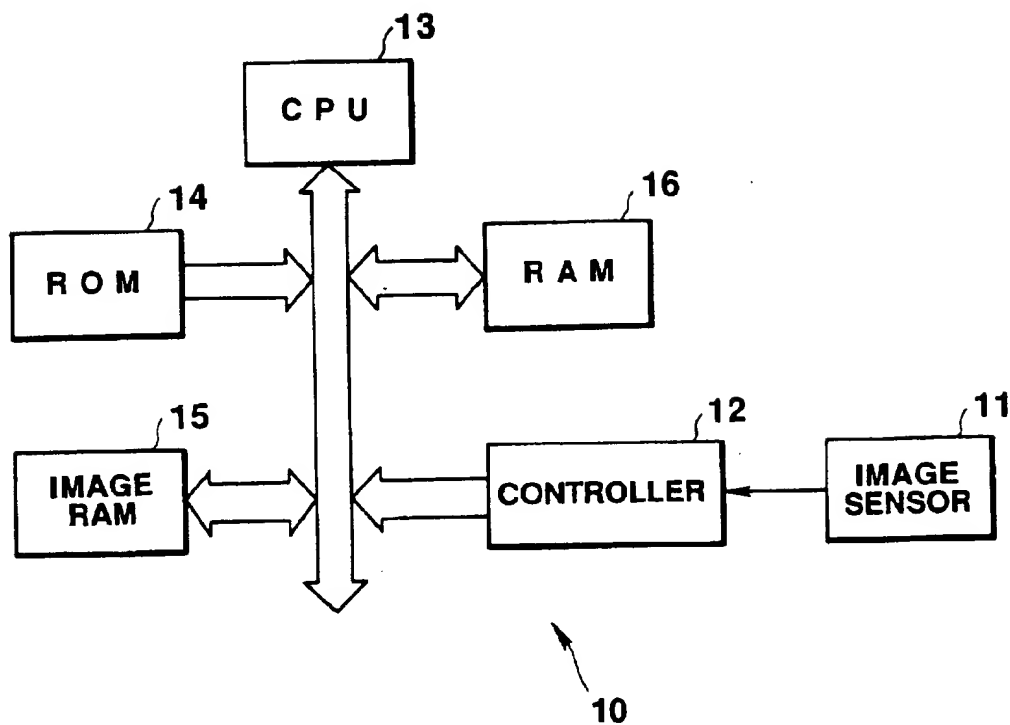


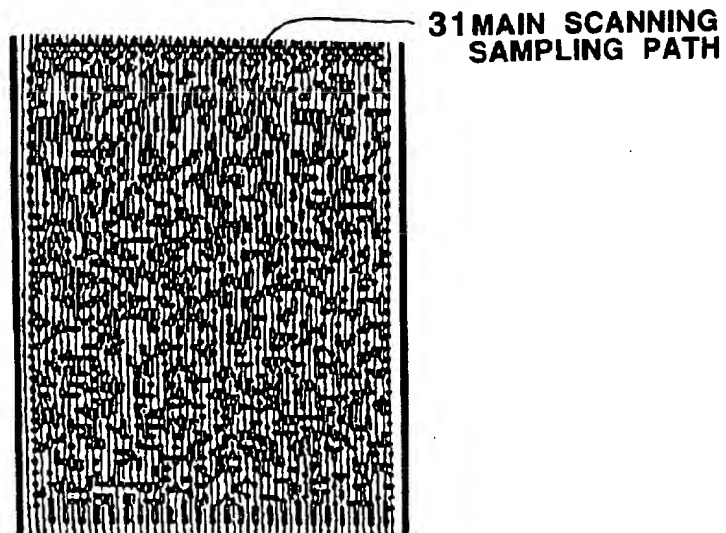
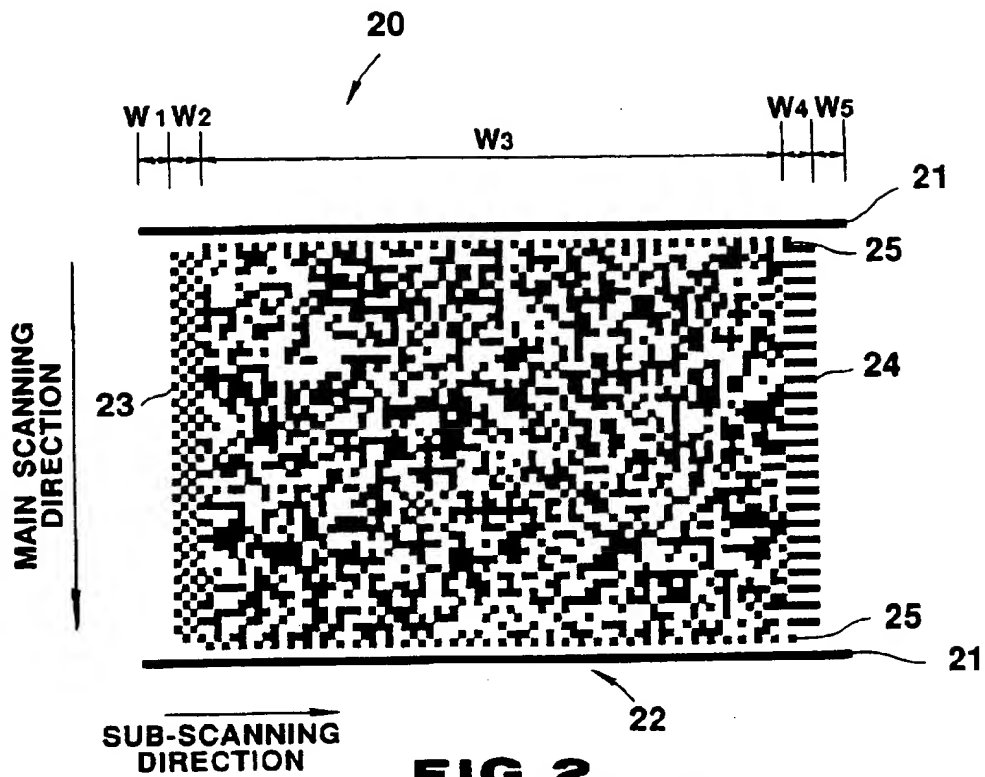
5,664,030

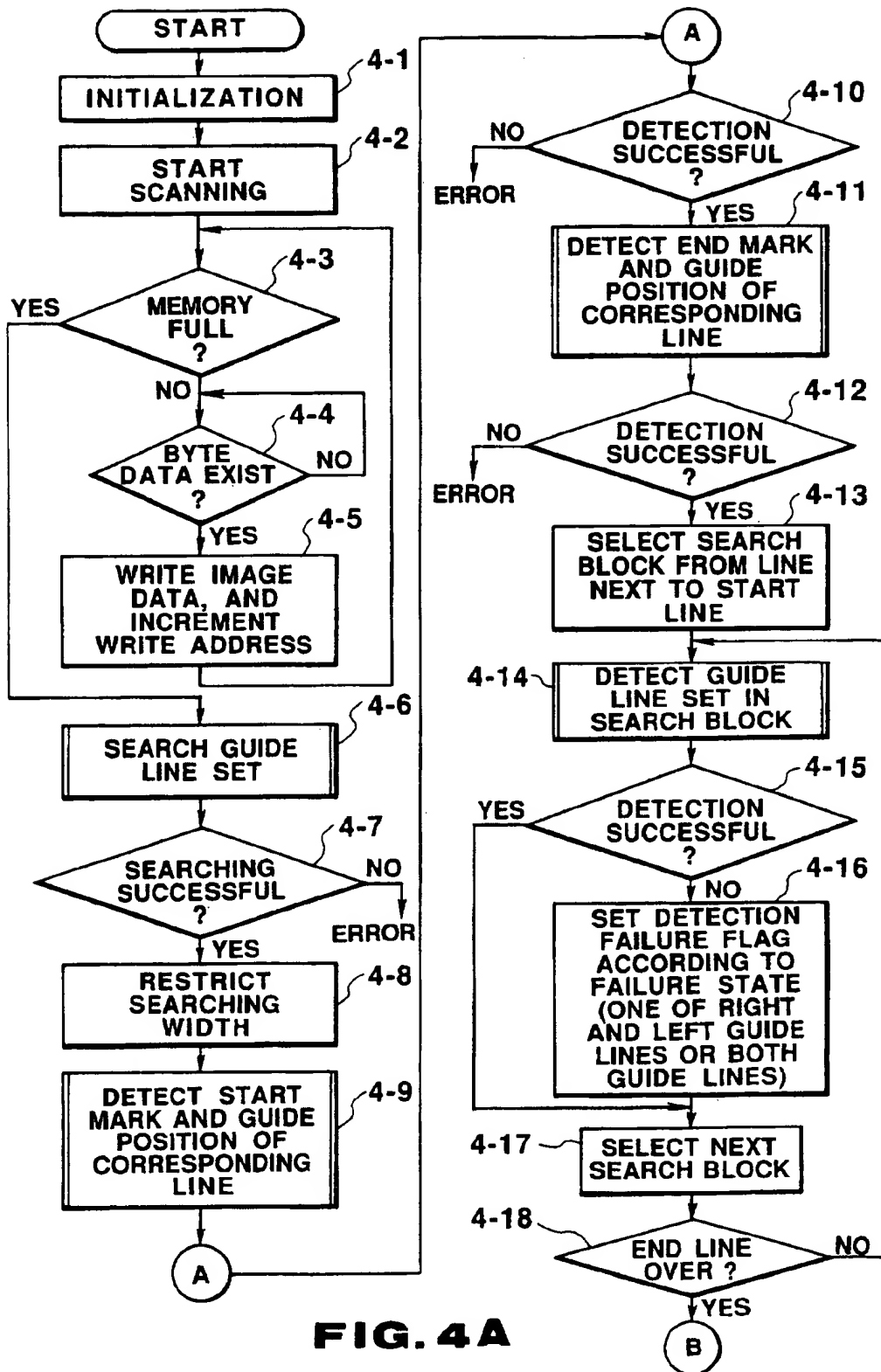
Page 2

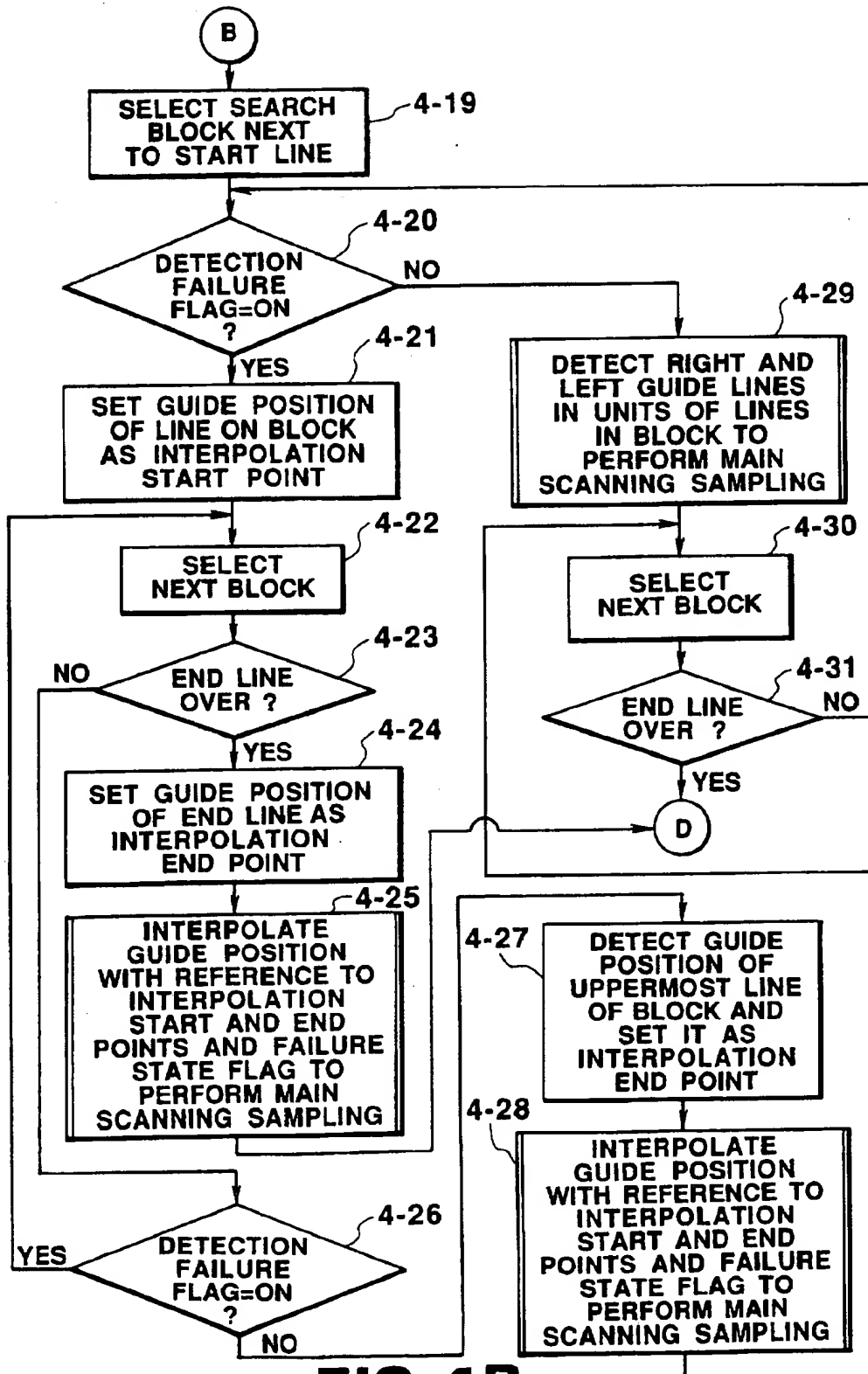
U.S. PATENT DOCUMENTS

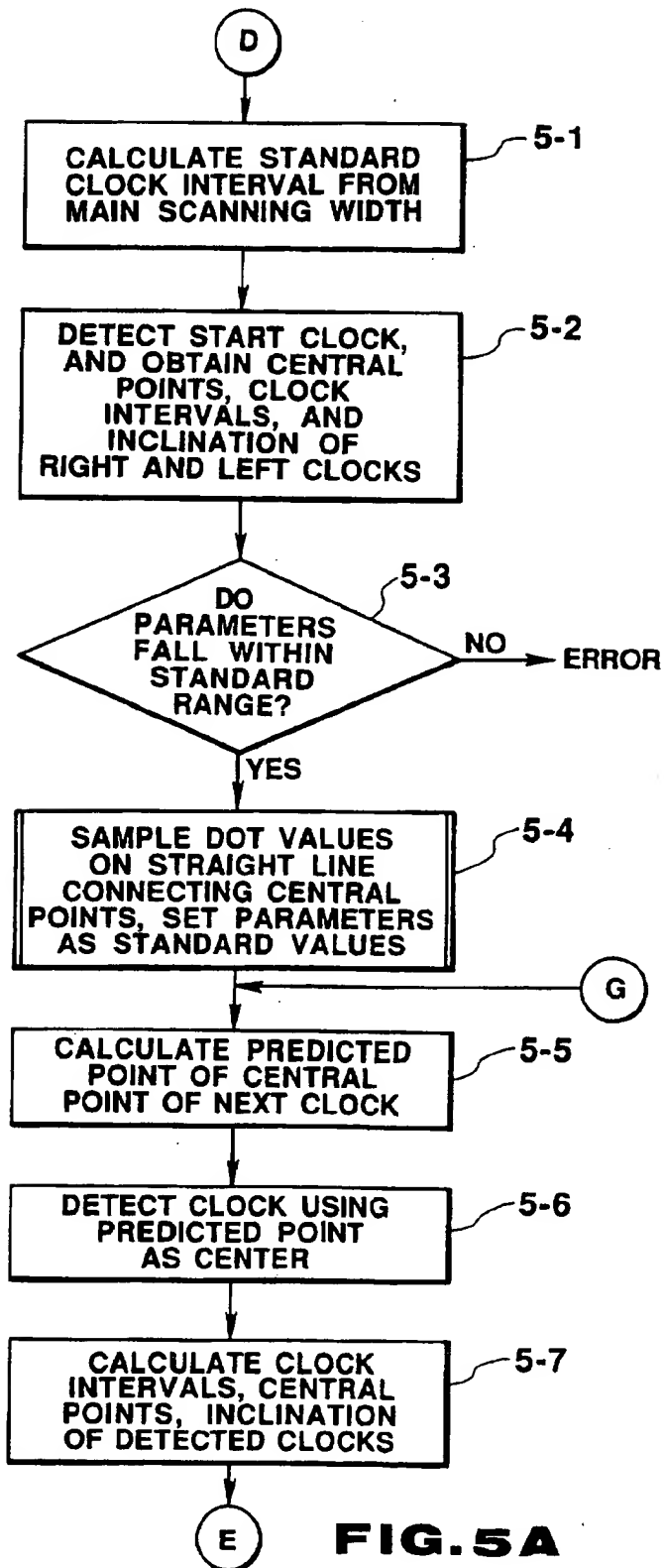
4,972,497	11/1990	Saito	382/56	5,038,393	8/1991	Nanba	382/61
4,974,260	11/1990	Rudak	382/57	5,053,609	10/1991	Priddy et al.	235/436
4,982,077	1/1991	Kawamura	235/494	5,124,536	6/1992	Priddy et al.	235/436
				5,126,542	6/1992	Priddy et al.	235/456
				5,304,786	4/1994	Pavlidis et al.	235/462

**FIG. 1**





**FIG. 4B**

**FIG. 5A**

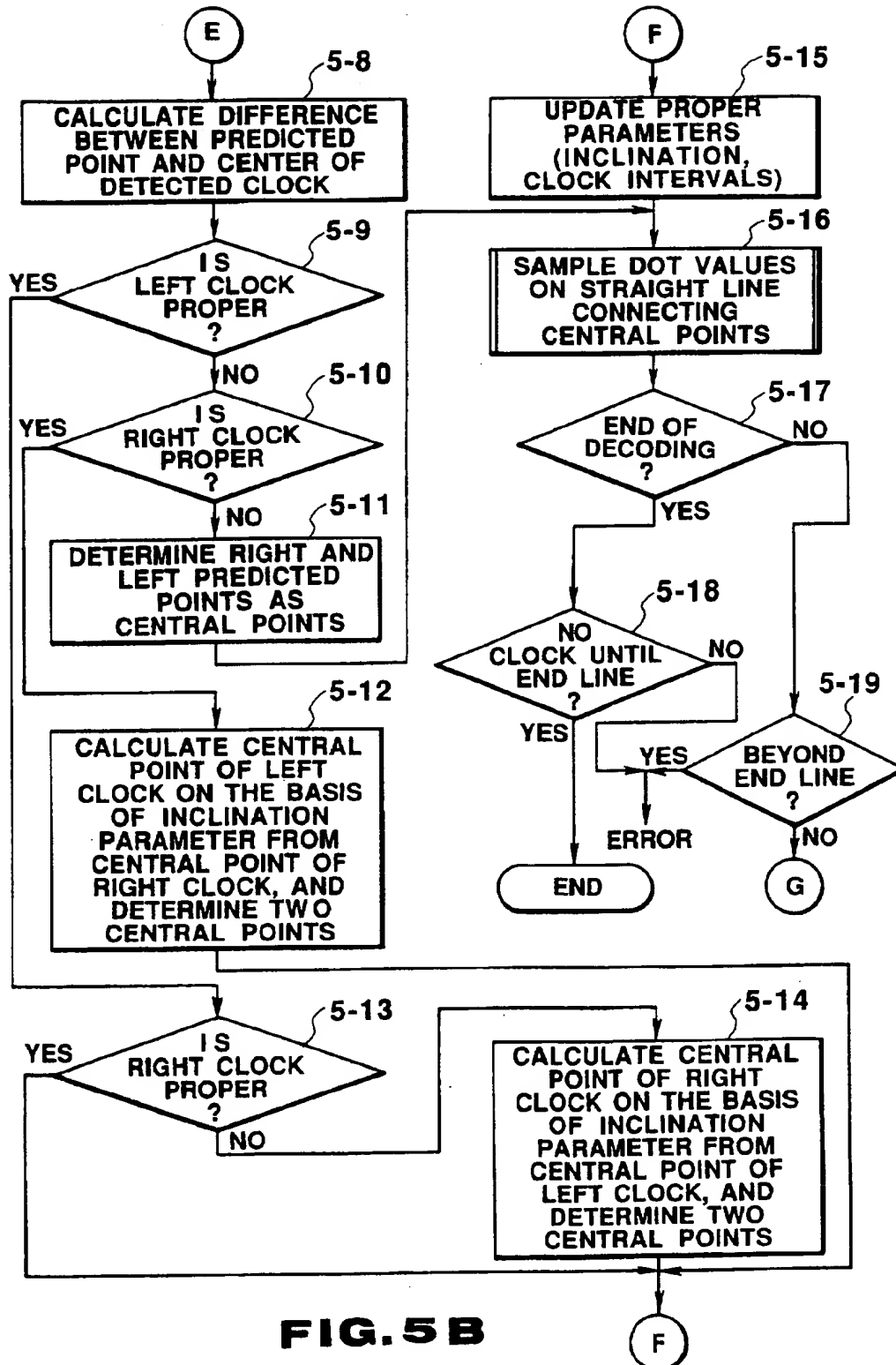
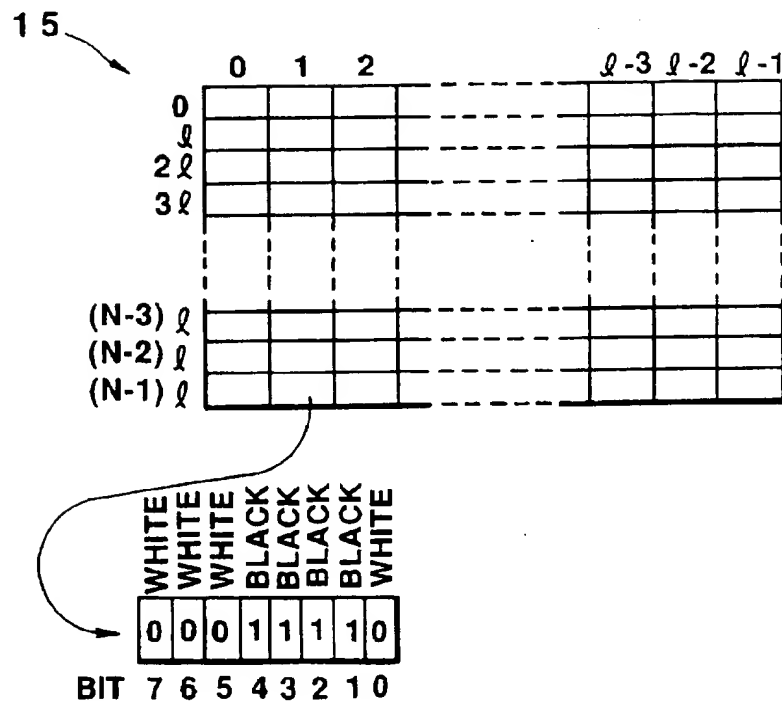
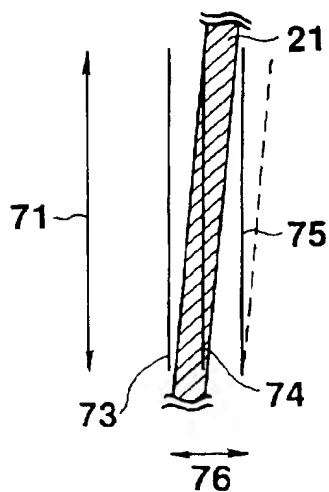
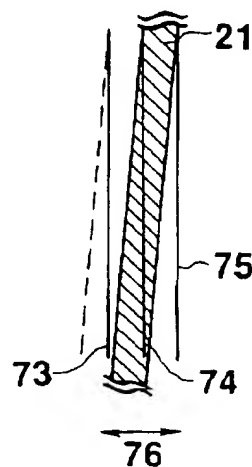
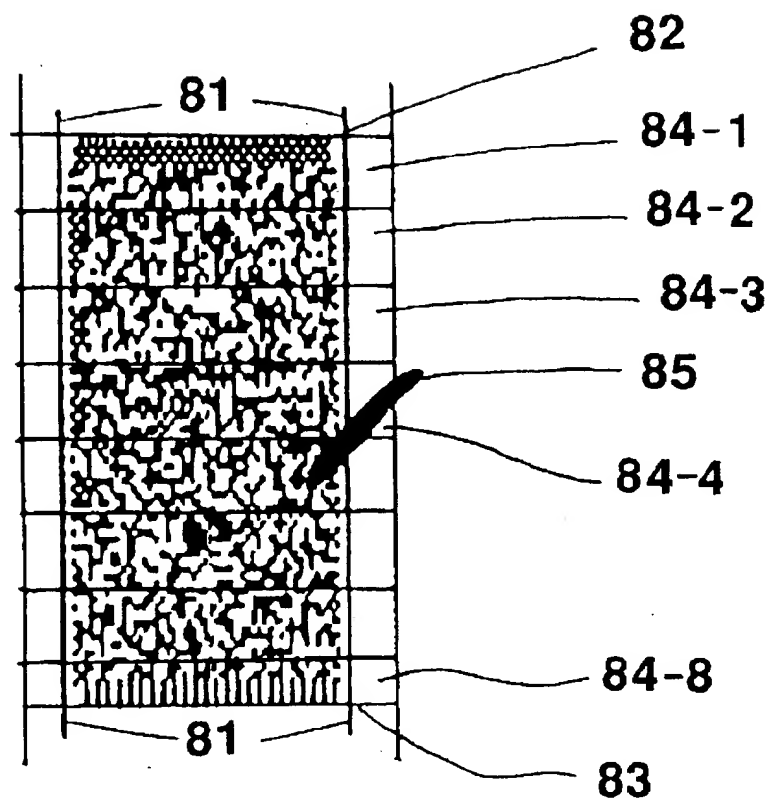
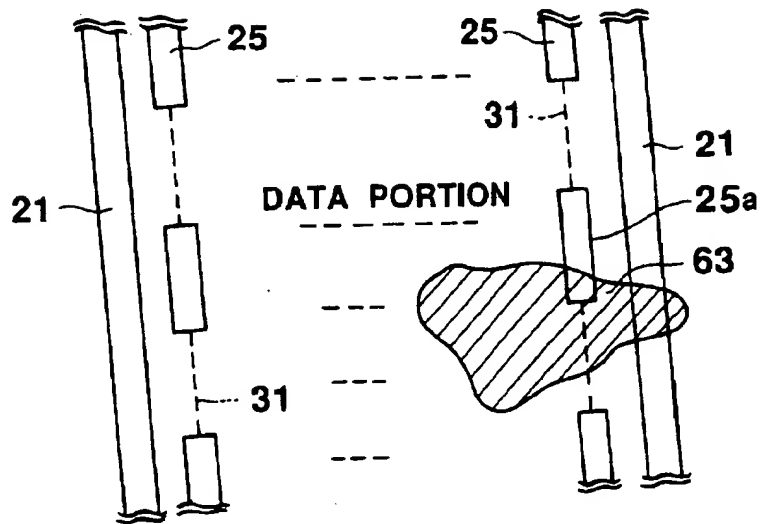


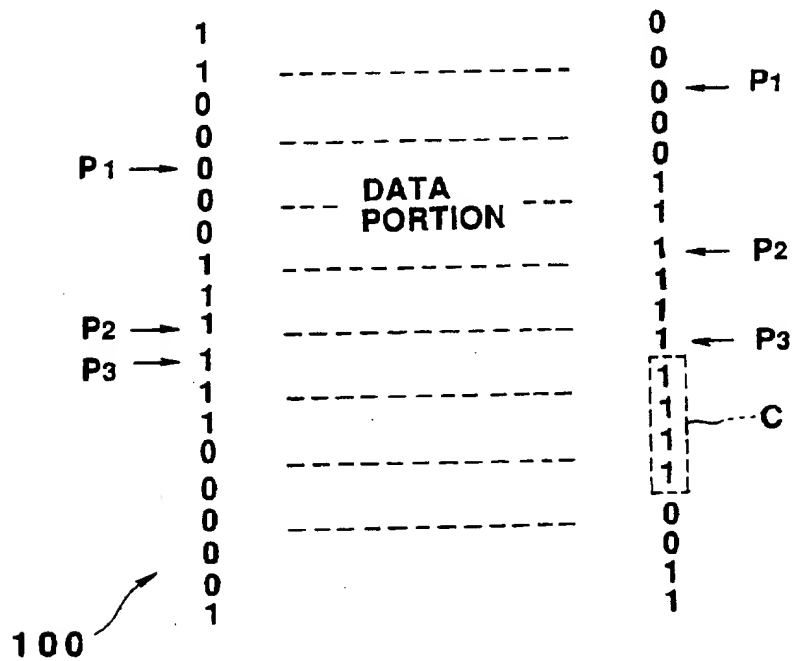
FIG. 5B

**FIG. 6****FIG. 7 (a)****FIG. 7 (b)**

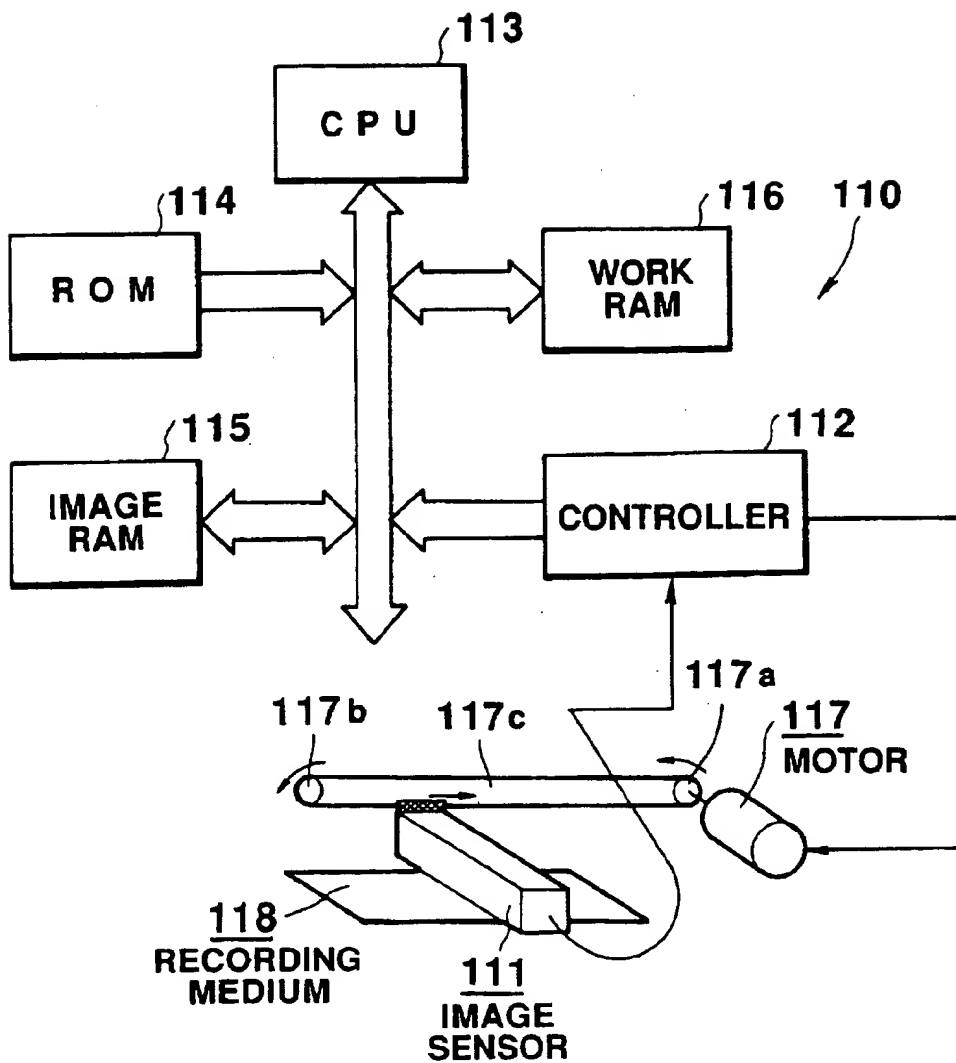
**FIG. 8**

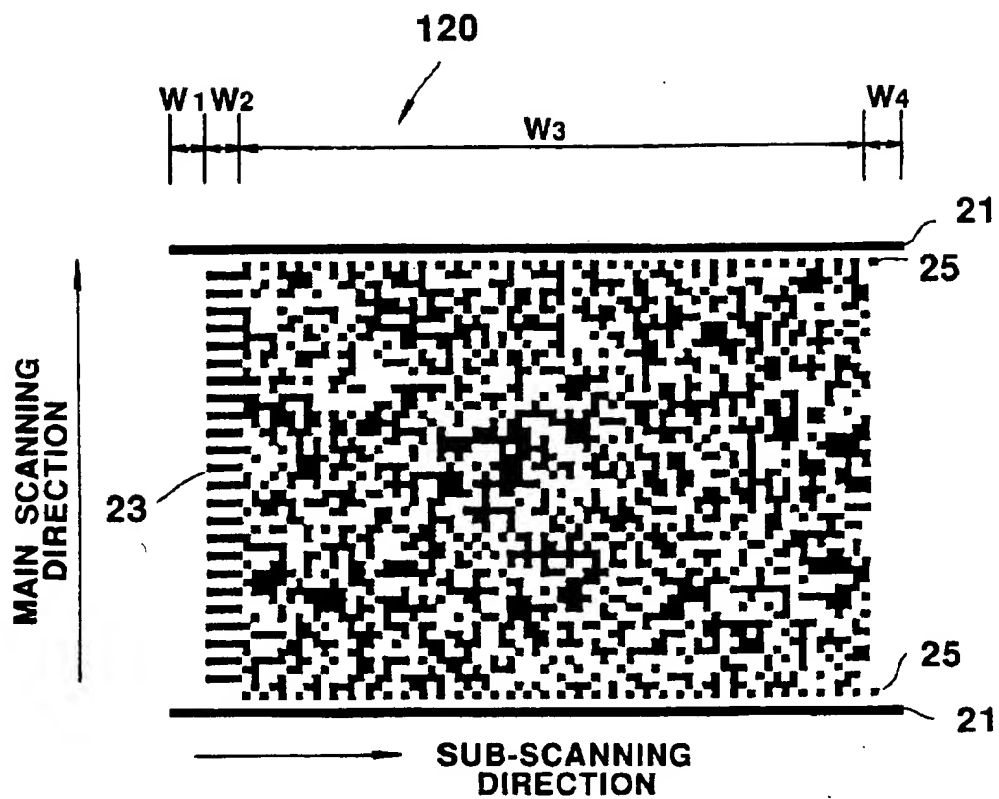


**FIG. 9**

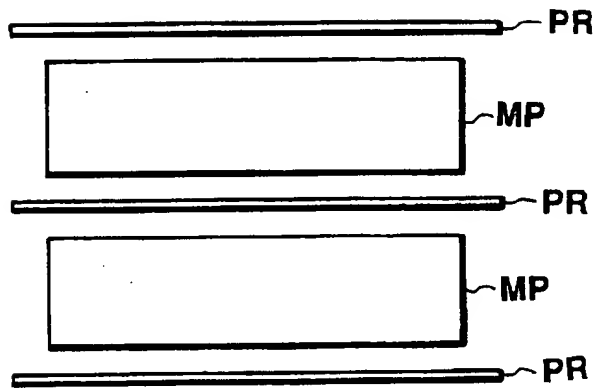


**FIG. 10**

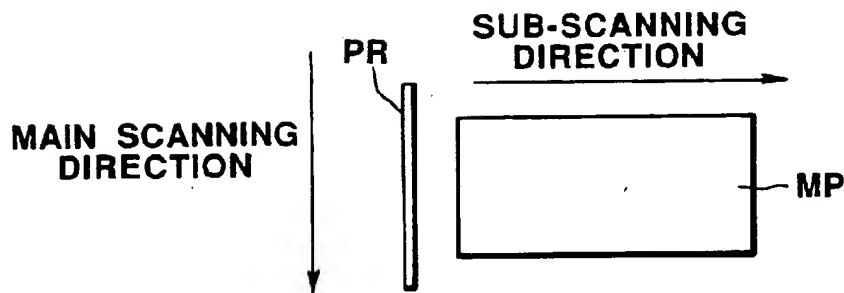
**FIG.11**



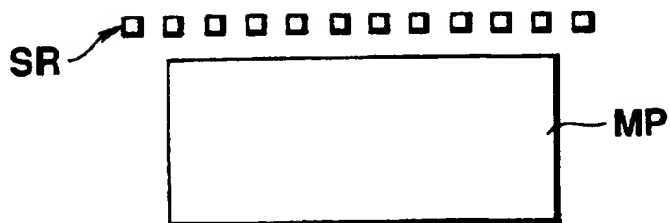
**FIG. 12**



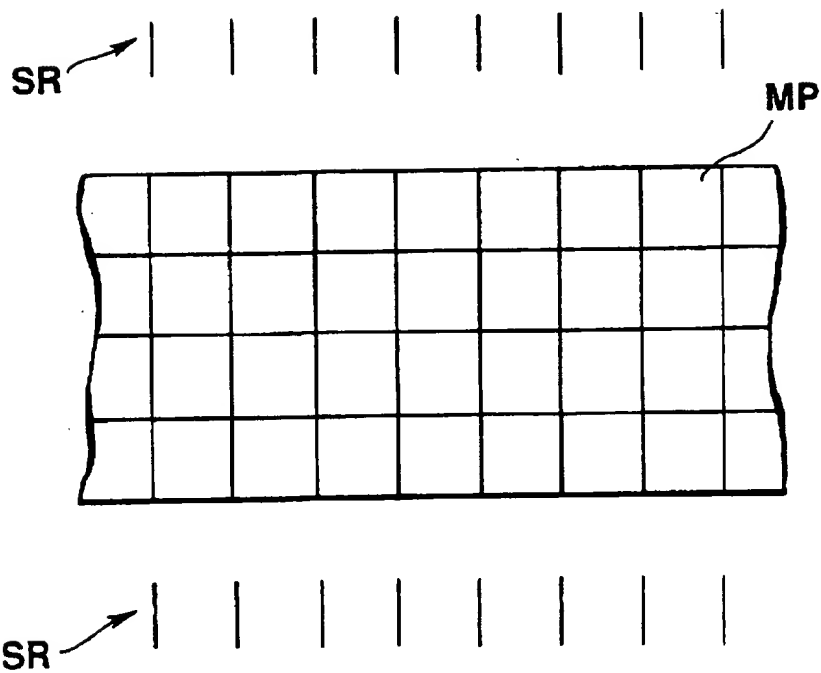
**FIG. 13**



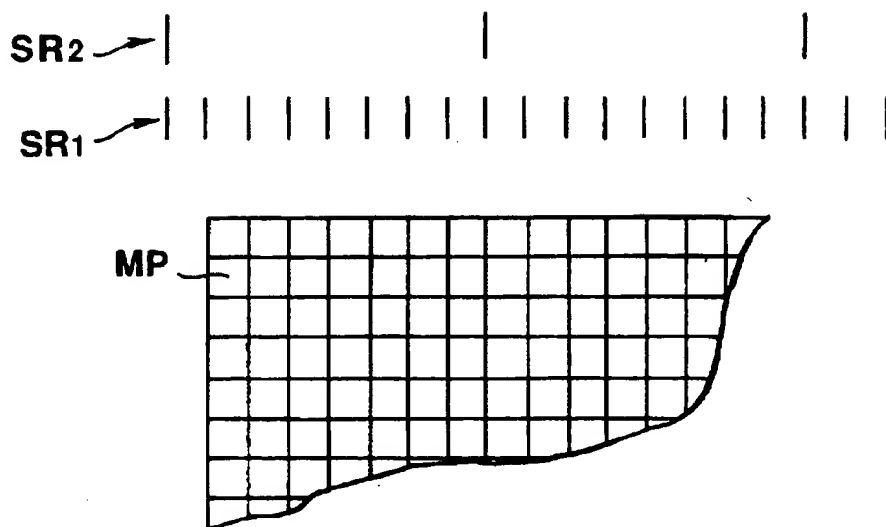
**FIG. 14**



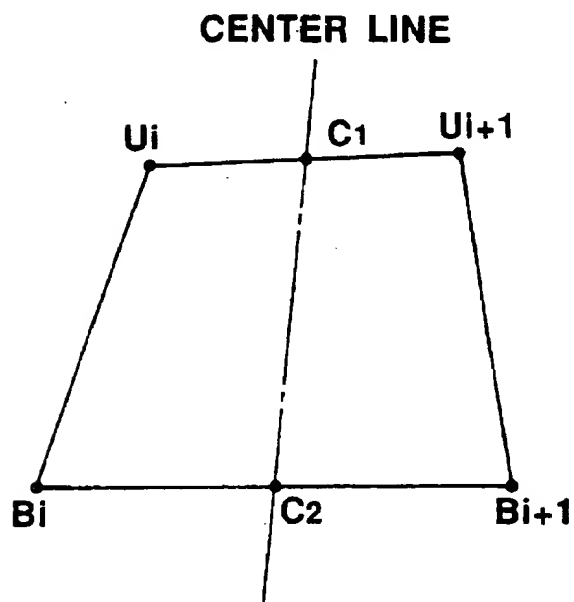
**FIG. 15**



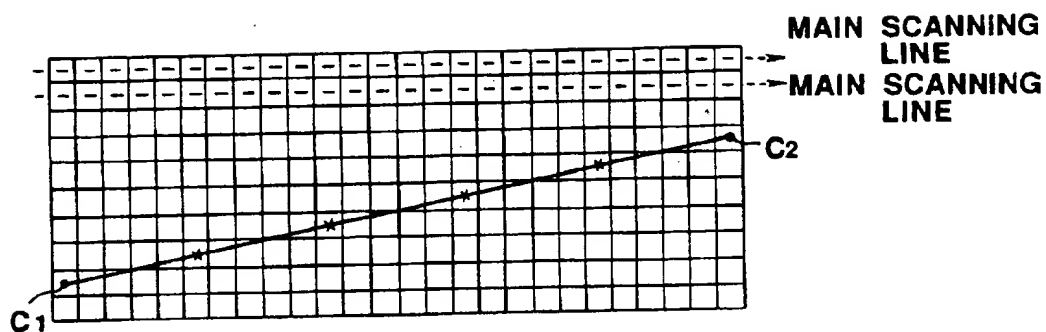
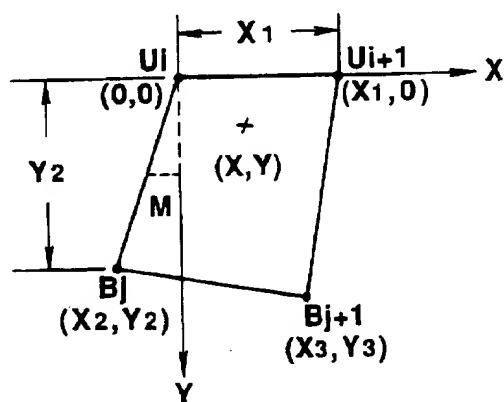
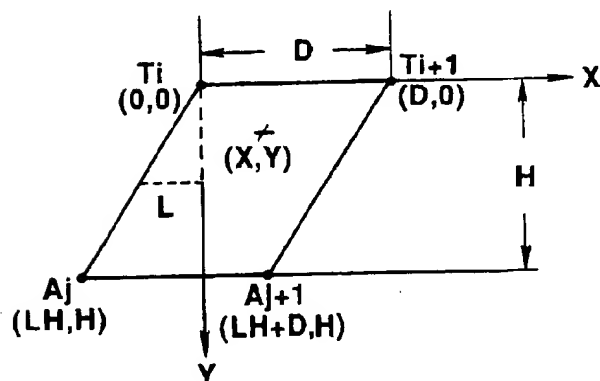
**FIG. 16**

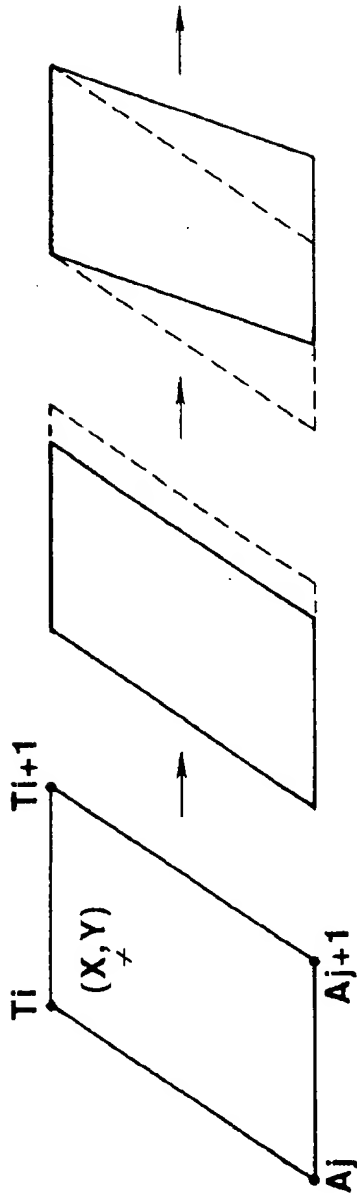


**FIG. 17**

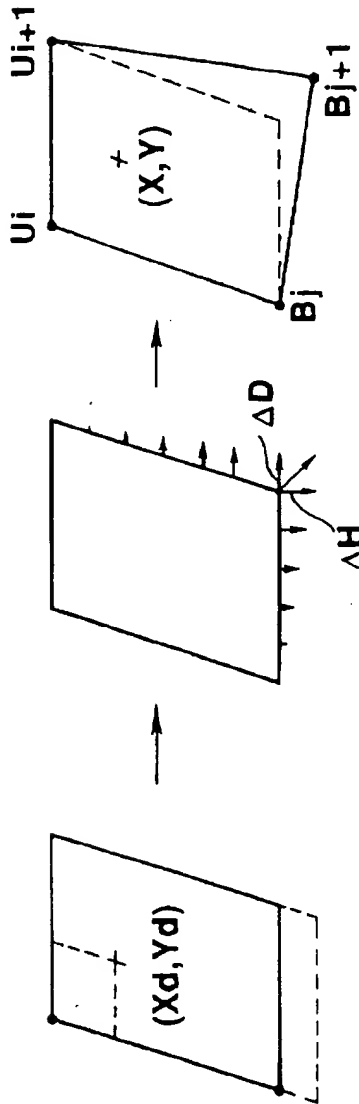


**FIG. 18**

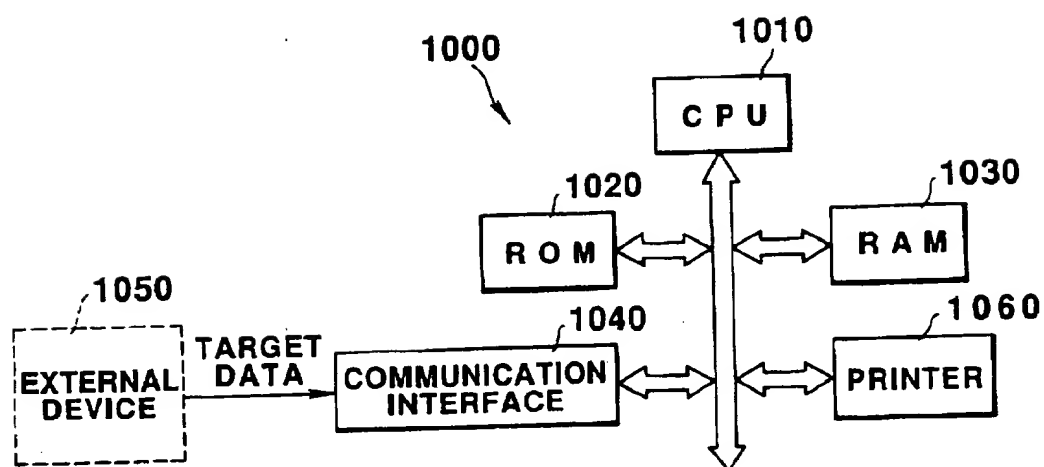
**FIG. 19****FIG. 20****FIG. 21**

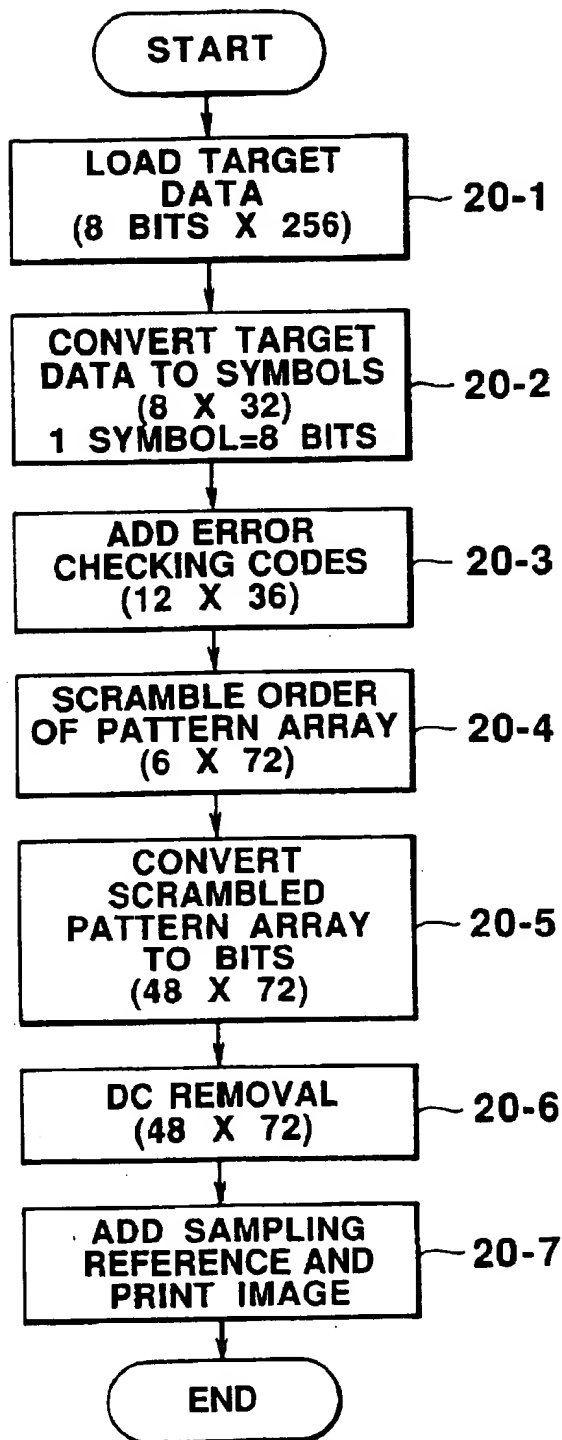


**FIG. 22(a) FIG. 22(b) FIG. 22(c)**



**FIG. 22(d) FIG. 22(e) FIG. 22(f)**

**FIG. 23**

**FIG. 24**

300

d0	d64	d128	d192	d256	d320	.	.	d1920	d1984
d1	d65	d129	d193	d257	d321	.	.	.	d1985
d2	d66	.	.	.	.	.	.	.	d1986
d3	.	.	.	.	.	.	.	.	d1987
d4	.	.	.	.	.	.	.	.	d1988
d5	.	.	.	.	.	.	.	.	.
d6	.	.	.	.	.	.	.	.	.
d7	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
d62	d126	d190	d254	d318	d382	.	.	.	d2046
d63	d127	d191	d255	d319	d383	.	.	d1983	d2047

FIG. 25

s0	s8	s16	s24	s32	s40	"	"	s240	s248
s1	s9	s17	s25	s33	s41	"	"	s241	s249
s2	s10	s18	s26	s34	s42	"	"	s242	s250
s3	s11	s19	s27	s35	s43	"	"	s243	s251
s4	s12	s20	s28	s36	s44	"	"	s244	s252
s5	s13	s21	s29	s37	s45	"	"	s245	s253
s6	s14	s22	s30	s38	s46	"	"	s246	s254
s7	s15	s23	s31	s39	s47	"	"	s247	s255

400

**FIG. 26**

s0	s8	s16	s24	s32	s40	"	s232	s240	s248	p0	p8	p16	p24
s1	s9	s17	s25	s33	s41	"	s233	s241	s249	p1	p9	p17	p25
s2	s10	s18	s26	s34	s42	"	s234	s242	s250	p2	p10	p18	p26
s3	s11	s19	s27	s35	s43	"	s235	s243	s251	p3	p11	p19	p27
s4	s12	s20	s28	s36	s44	"	s236	s244	s252	p4	p12	p20	p28
s5	s13	s21	s29	s37	s45	"	s237	s245	s253	p5	p13	p21	p29
s6	s14	s22	s30	s38	s46	"	s238	s246	s254	p6	p14	p22	p30
s7	s15	s23	s31	s39	s47	"	s239	s247	s255	p7	p15	p23	p31
q0	q4	q8	q12	q16	q20	"	q116	q120	q124	q128	q132	q136	q140
q1	q5	q9	q13	q17	q21	"	q117	q121	q125	q129	q133	q137	q141
s2	q6	q10	q14	q18	q22	"	q118	q122	q126	q130	q134	q138	q142
s3	q7	q11	q15	q19	q23	"	q119	q123	q127	q131	q135	q139	q143

500

FIG. 27

s0	s8	s16	s24	s32	s40	"	s232	s240	s248	p0	p8	p16	p24
s1	s9	s17	s25	s33	s41	"	s233	s241	s249	p1	p9	p17	p25
q0	q4	q8	q12	q16	q20	"	q116	q120	q124	q128	q132	q136	q140
q1	q5	q9	q13	q17	q21	"	q117	q121	q125	q129	q133	q137	q141
s2	s10	s18	s26	s34	s42	"	s234	s242	s250	p2	p10	p18	p26
s3	s11	s19	s27	s35	s43	"	s235	s243	s251	p3	p11	p19	p27
s4	s12	s20	s28	s36	s44	"	s236	s244	s252	p4	p12	p20	p28
s5	s13	s21	s29	s37	s45	"	s237	s245	s253	p5	p13	p21	p29
q2	q6	q10	q14	q18	q22	"	q118	q122	q126	q130	q134	q138	q142
q3	q7	q11	q15	q19	q23	"	q119	q123	q127	q131	q135	q139	q143
s6	s14	s22	s30	s38	s46	"	s238	s246	s254	p6	p14	p22	p30
s7	s15	s23	s31	s39	s47	"	s239	s247	s255	p7	p15	p23	p31

600

FIG. 28A

s	s	p
2*16	2*16	2*4
q	q	q
2*16	2*16	2*4
s	s	p
2*16	2*16	2*4
s	s	p
2*16	2*16	2*4
q	q	q
2*16	2*16	2*4
s	s	p
2*16	2*16	2*4

600S

FIG. 28B

700

	0	1	31	32	35	36	67	68	71				
0	s0	s8	"	s248	p0	"	p24	s4	"	s252	p4	"	p28
1	s1	s9	"	s249	p1	"	p25	s5	"	s253	p5	"	p29
2	q0	q4	"	q124	q128	"	q140	q2	"	q126	q130	"	q142
3	q1	q5	"	q125	q129	"	q141	q3	"	q127	q131	"	q143
4	s2	s10	"	s250	p2	"	p26	s6	"	s254	p6	"	p30
5	s3	s11	"	s251	p3	"	p27	s7	"	s255	p7	"	p31

**FIG. 29A**

700S

s	2*16	s	2*16	p	2*4	s	2*16	s	2*16	p	2*4
q	2*16	q	2*16	q	2*4	q	2*16	q	2*16	q	2*4
s	2*16	s	2*16	p	2*4	s	2*16	s	2*16	p	2*4

**FIG. 29B**

0	15 16			19	20	51	52	55 56			71				
0	s128	"	s248	p0	"	p24	s4	"	s252	p4	"	p28	s0	"	s120
1	s129	"	s249	p1	"	p25	s5	"	s253	p5	"	p29	s1	"	s121
2	q64	"	q124	q128	"	q140	q2	"	q126	q130	"	q142	q0	"	q60
3	q65	"	q125	q129	"	q141	q3	"	q127	q131	"	q143	q1	"	q61
4	s130	"	s250	p2	"	p26	s6	"	s254	p6	"	p30	s2	"	s122
5	s131	"	s251	p3	"	p27	s7	"	s255	p7	"	p31	s3	"	s123

**FIG. 30A**800

s	2*16	p	2*4	s	2*16	s	2*16	p	2*4	s	2*16
q	2*16	q	2*4	q	2*16	q	2*16	q	2*4	q	2*16
s	2*16	p	2*4	s	2*16	s	2*16	p	2*4	s	2*16

**FIG. 30B**800S

900 ↗

b0	b48	b96	b144	"	b3360	b3408
b1	b49	.	.	"	.	b3409
b2	b50	.	.	"	.	b3410
b3	.	.	.	"	.	b3411
b4	.	.	.	"	.	.
b5	.	.	.	"	.	.
b6	.	.	.	"	.	.
b7	.	.	.	"	.	.
.	.	.	.	"	.	.
.	.	.	.	"	.	.
b46	b94	.	.	"	.	b3454
b47	b95	b143	b191	"	b3407	b3455

FIG. 31

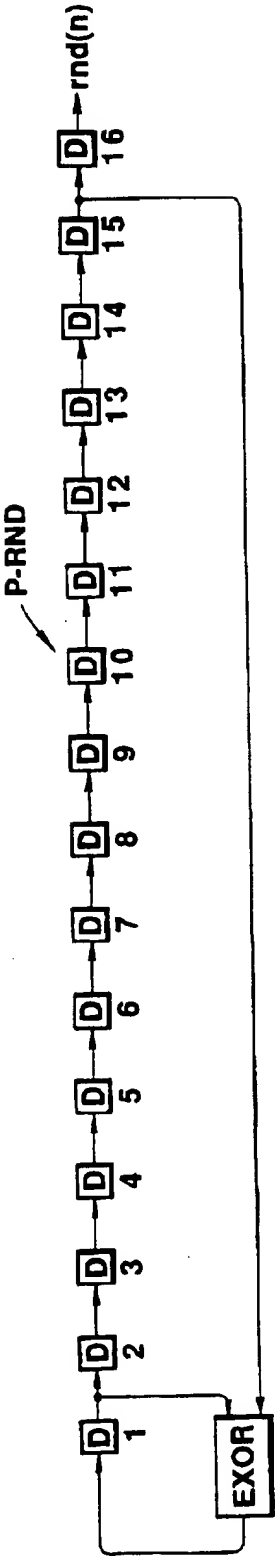
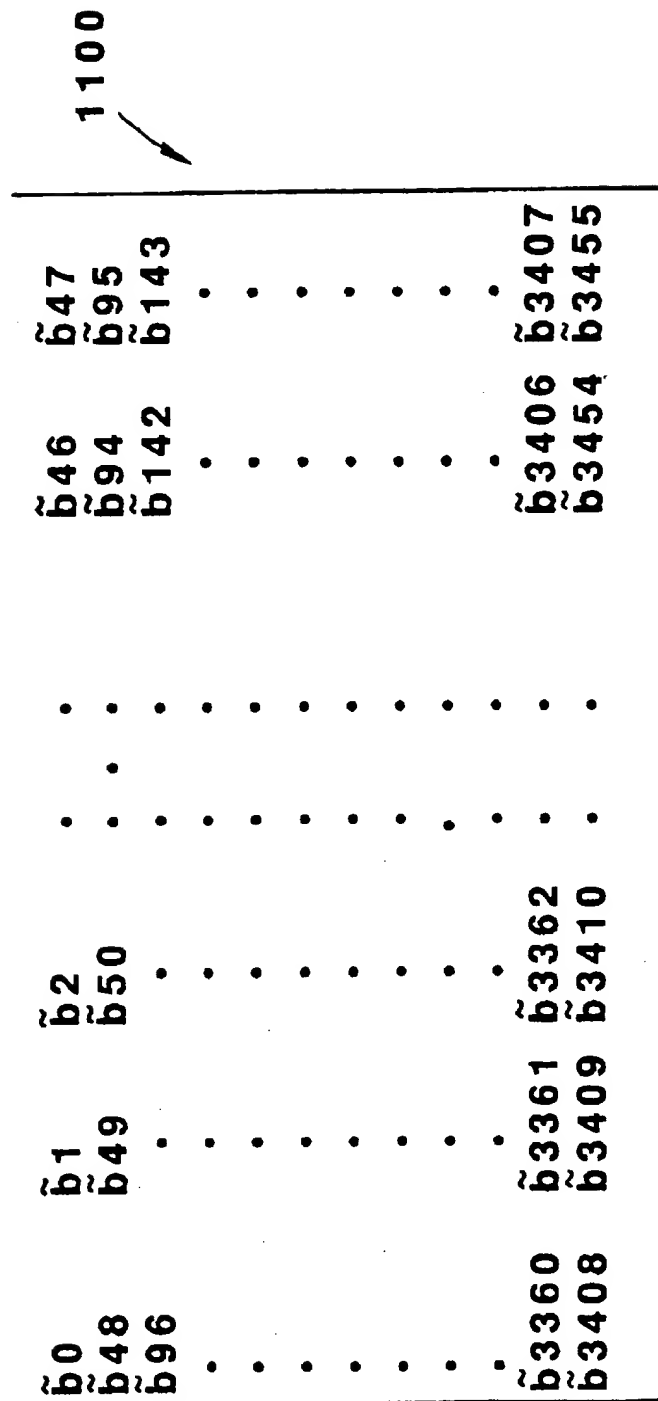
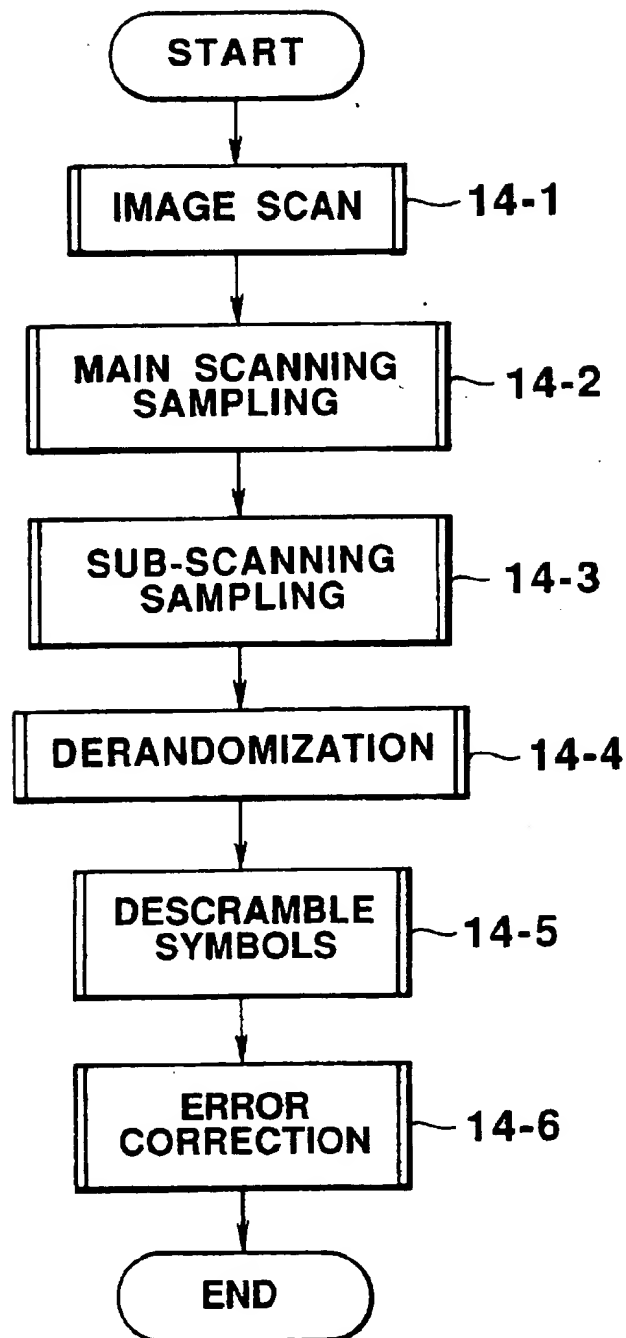


FIG. 32



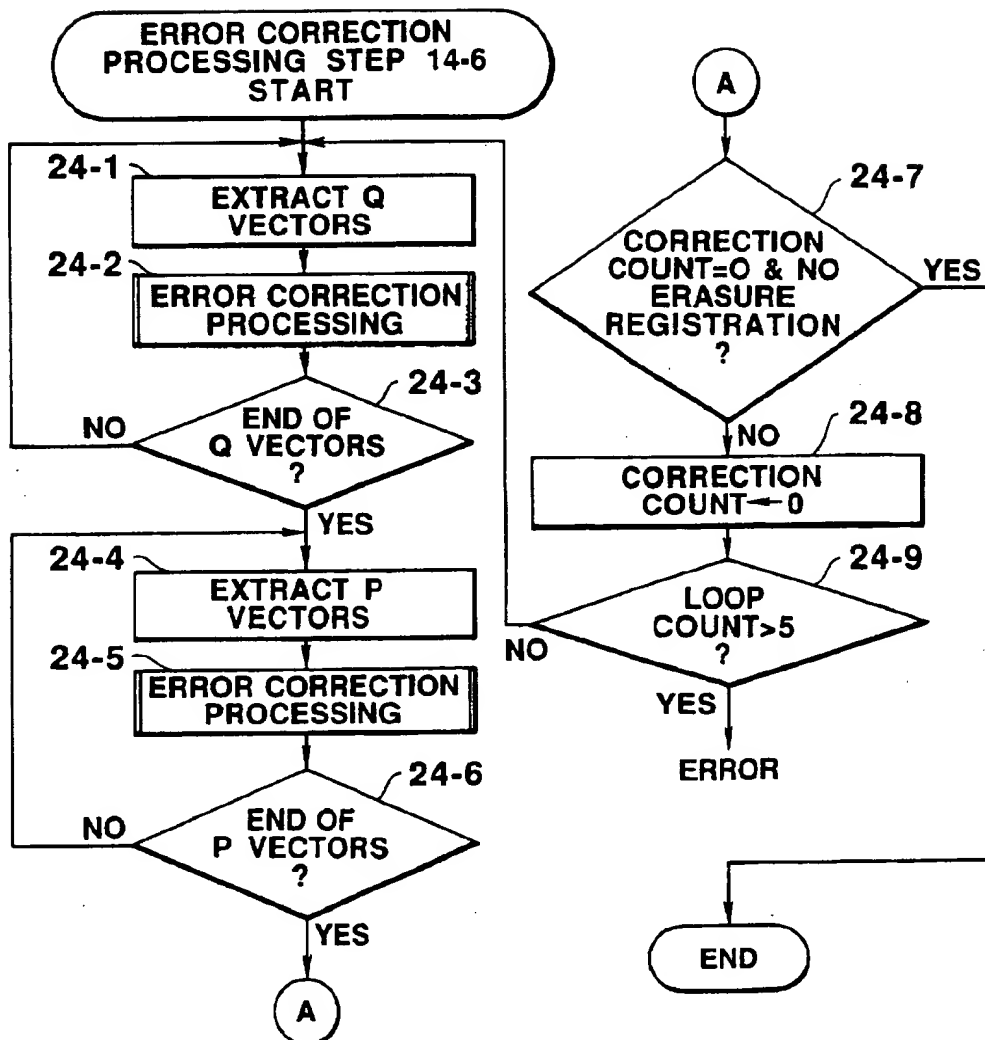
**FIG. 3**

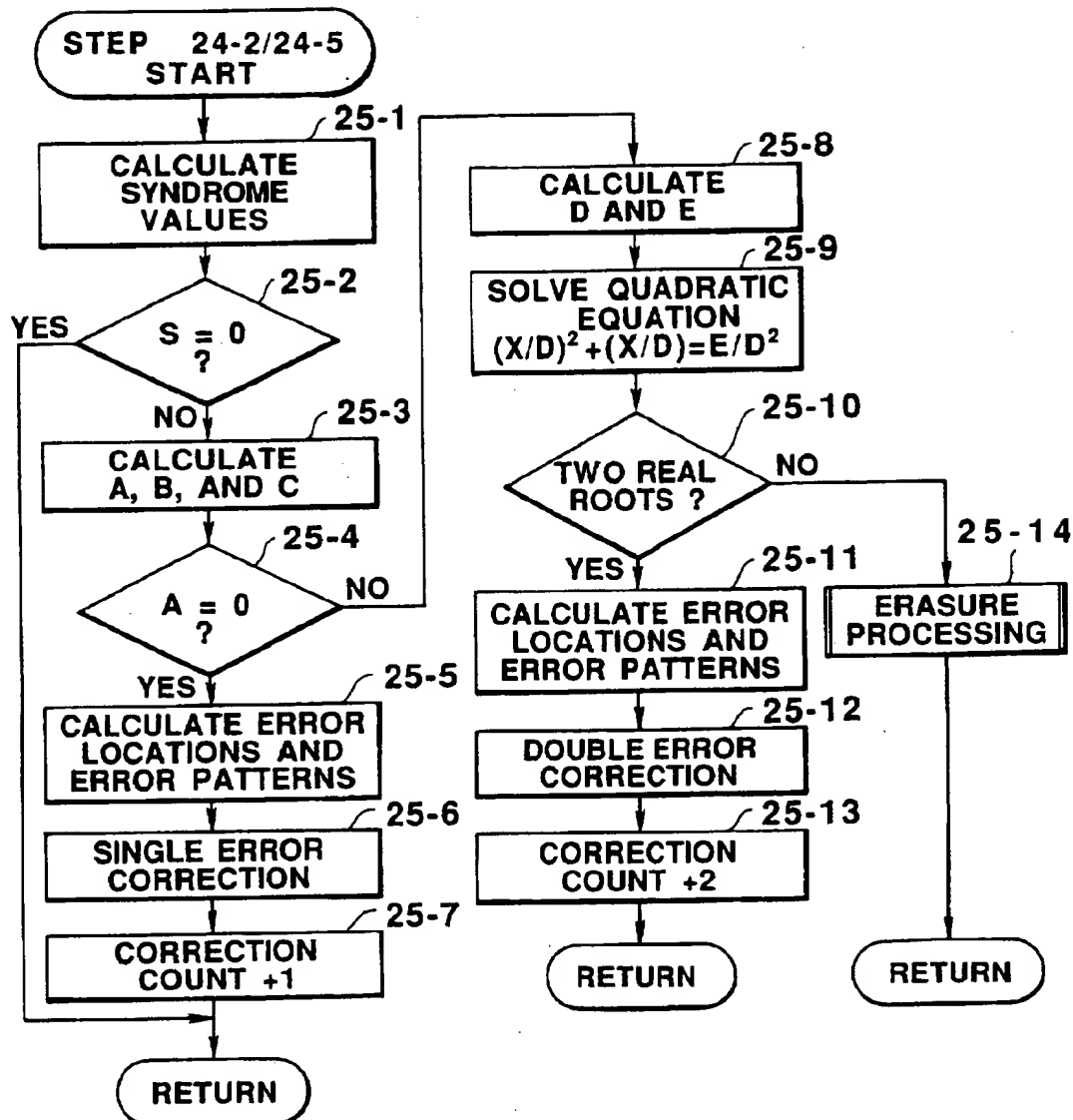
**FIG. 34**

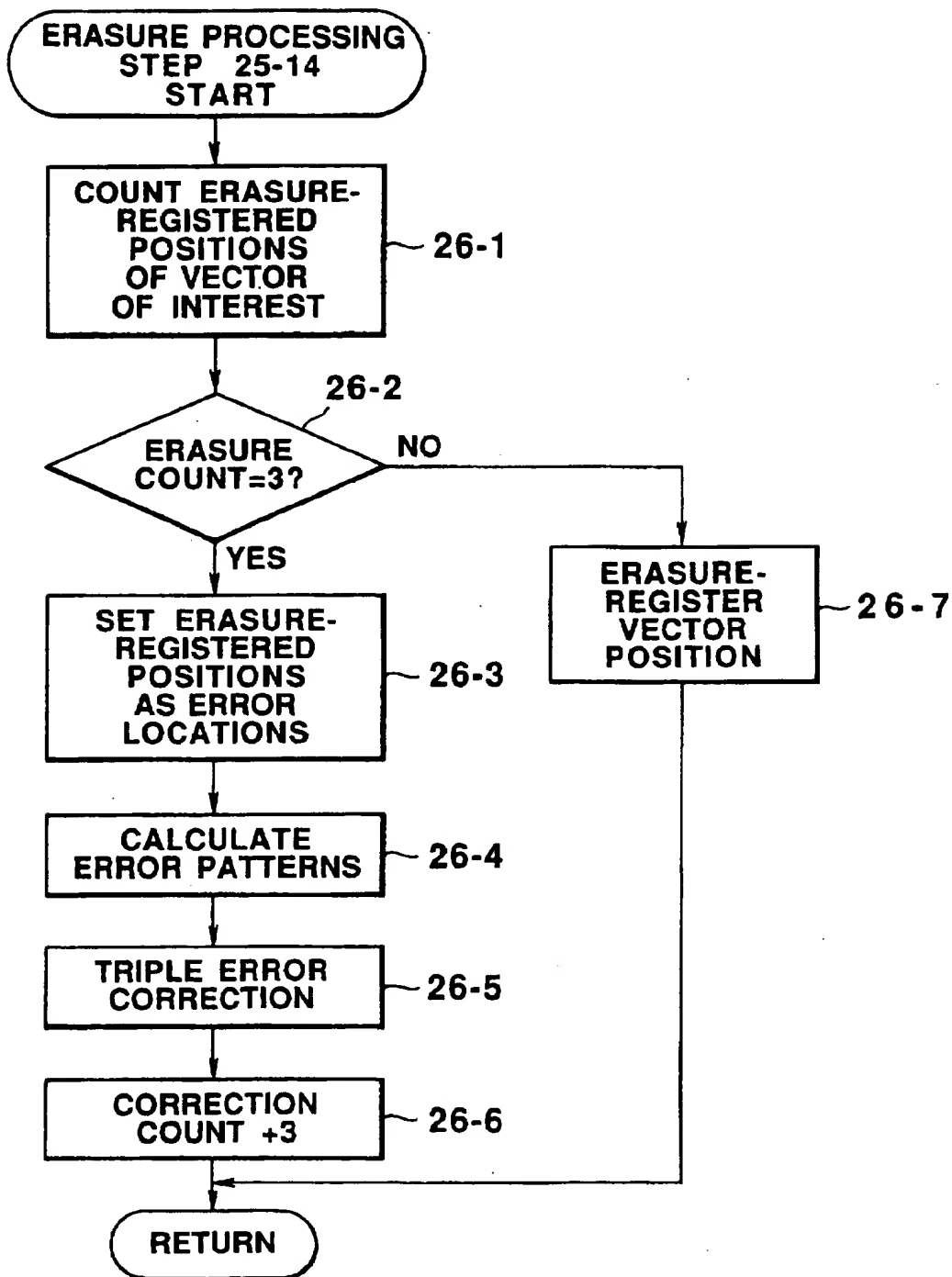
500R

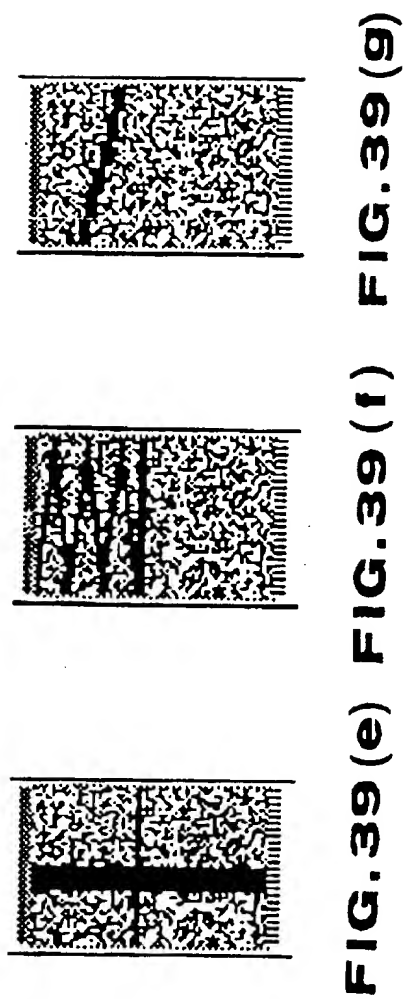
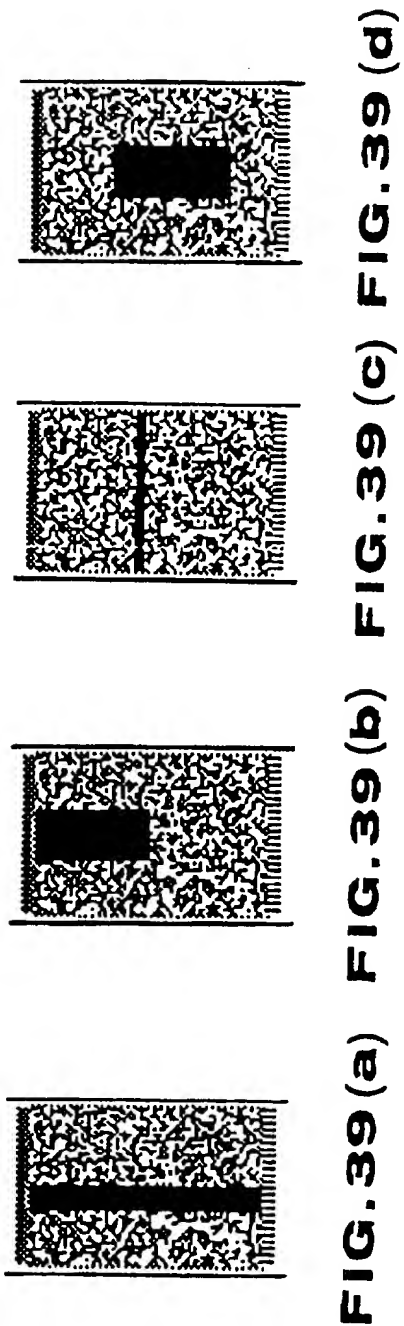
S0	S8	S16	S24	S32	S40	"	S232	S240	S248	P0	P8	P16	P24	P0
S1	S9	S17	S25	S33	S41	"	S233	S241	S249	P1	P9	P17	P25	
S2	S10	S18	S26	S34	S42	"	S234	S242	S250	P2	P10	P18	P26	
S3	S11	S19	S27	S35	S43	"	S235	S243	S251	P3	P11	P19	P27	
S4	S12	S20	S28	S36	S44	"	S236	S244	S252	P4	P12	P20	P28	
S5	S13	S21	S29	S37	S45	"	S237	S245	S253	P5	P13	P21	P29	
S6	S14	S22	S30	S38	S46	"	S238	S246	S254	P6	P14	P22	P30	
S7	S15	S23	S31	S39	S47	"	S239	S247	S255	P7	P15	P23	P31	P7
Q0	Q4	Q8	Q12	Q16	Q20	"	Q116	Q120	Q124	Q128	Q132	Q136	Q140	
Q1	Q5	Q9	Q13	Q17	Q21	"	Q117	Q121	Q125	Q129	Q133	Q137	Q141	
Q2	Q6	Q10	Q14	Q18	Q22	"	Q118	Q122	Q126	Q130	Q134	Q138	Q142	
Q3	Q7	Q11	Q15	Q19	Q23	"	Q119	Q123	Q127	Q131	Q135	Q139	Q143	Q35

FIG. 35

**FIG. 36**

**FIG. 37**

**FIG. 38**



# METHOD AND APPARATUS FOR RECORDING/REPRODUCING MESH PATTERN DATA

This is a division of application Ser. No. 07/869,012 filed Apr. 14, 1992 now U.S. Pat. No. 5,454,054, which is a Division of Ser. No. 07/530,630 filed May 30, 1990 (now U.S. Pat. No. 5,153,928).

## BACKGROUND OF THE INVENTION

### 1. Field of the Invention

The present invention relates to a data recording/reproduction system for encoding binary data to record same as an image, and for reading a recorded image and decoding the read image into binary data.

### 2. Description of the Related Art

A bar code technique is known as a technique for reading an encoded image from a recording medium (recording sheet) to reproduce binary data.

For example, U.S. Pat. Nos. 4,422,361, 4,437,378, and 4,464,966 disclose this technique. However, it is difficult to increase a recording density of a bar code in terms of its structure, and the bar code is not suitable for handling a large volume of data.

Unexamined Japanese Patent Application Publication No. 53-73026 discloses a technique wherein an image in which some meshes of an  $i \times j$  (e.g.,  $3 \times 3$ ) matrix are set to be black meshes, and the remaining meshes are set to be white meshes; is read to recognize a black-and-white mesh pattern. The use of the matrix encoded image can easily increase a data volume by increasing the number of meshes included in the matrix. Another technique using a similar matrix image is also described in Magazine I/O, May, 1988, pp. 121 to 125, "Damp List Read by Image Scanner". U.S. patent application Ser. No. 07/389,287 (filing date: Aug. 3, 1989, inventors: Morikawa et al. now U.S. Pat. No. 5,042,079, assigned to the present assignee proposes the following technique. That is, a mesh pattern obtained by encoding data by black-and-white meshes selectively formed on a mesh matrix consisting of a large number of meshes is used as a data body of an encoded image in place of a bar code, and such an encoded image is read and decoded.

According to the invention disclosed in this U.S. Ser. No. 07/389,287, since data bits are expressed by a black-and-white pattern of meshes which are two-dimensionally arrayed at small intervals, a recording density can be greatly improved.

However, when an encoded image itself on a recording medium is stained, e.g., contaminated, and a scanning reference pattern is partially deformed, omitted, or damaged thereby, the above-mentioned decoding method may erroneously recognize the contamination as a segment of the scanning reference pattern, and position errors of meshes of the mesh pattern on image data which are determined on the basis of the recognition result occur, resulting in a wrong black-and-white recognition result. For example, if a certain white clock mesh (hereinafter referred to as a clock) of a synchronous pattern as a sub-scanning reference pattern is stained in black and becomes continuous with black clocks before and after this white clock, when clocks of the synchronous pattern are detected on image data, the white clock is not read (i.e., three, i.e., black, white, and black clocks are detected as one black clock), and this error causes a decoding error of all other clocks. In this state, even if an error is detected later in error correction processing using a

checking code (redundant code included in a mesh pattern), correction is impossible to achieve since the error is considerably beyond a correction capability.

A position error of a guide line as a main scanning reference caused by contamination influences a position as a central position of meshes in the vertical direction thereof obtained by equally dividing a line connecting corresponding positions of two guide lines on a scanning line image. As a result, a noncorrectable sampling error occurs later.

A recognition error of image data may occur due to various other causes. For example, the causes include a defect (caused by poor printing precision, a variation in density, or the like) of an encoded image on a recording medium, problems of an image sensor (an image resolution of the image sensor, an image distortion in image processing by the image sensor (distortion of a lens, a change in scanning speed or direction), luminance stability of a light source, a threshold level for binarizing analog image data, and photoelectric conversion characteristics of an image sensor such as a CCD), and the like.

Most of such errors are caused by the fact that a mesh pattern is adopted to increase a recording density on a recording medium. In other words, most of errors are related to the fact that an encoded image on a recording medium is constituted by a two-dimensional array of image elements encoding information units in order to effectively use a two-dimensional space on the recording medium.

## SUMMARY OF THE INVENTION

It is therefore an object of the present invention to provide a data reading method and apparatus, which can detect correct positions of scanning reference patterns on image data even when scanning reference patterns of an encoded image having a mesh pattern as a data body are partially destroyed by, e.g., contamination on a recording medium.

It is another object of the present invention to provide a recording medium for recording an encoded image, which can effectively solve the above-mentioned error problems while maintaining a relatively high recording density, to provide a data recording method and apparatus for recording an encoded image on the recording medium, and to provide a data reproduction method and apparatus capable of reproducing desired data free from an error or at a low error rate from the encoded image.

According to one aspect of the present invention, image data representing an image is read from a recording medium which records the image including a mesh pattern and a main scanning reference pattern, and the read image data is divided into a plurality of partial image data segments. For each partial image data segment, a feature pixel group featuring a piece of the main scanning reference pattern included in each partial image data segment is searched from the partial image data piece. For the partial image data segment from which no feature pixel group is detected upon searching, the position of the piece of the main scanning reference pattern in that partial image data segment is determined from the position of a piece of the main scanning reference pattern represented by the feature pixel group detected upon searching of another partial image data segment. Thus, position data of pieces of the main scanning reference pattern in the image data are generated, and black and white meshes of the mesh pattern of the image data is recognized on the basis of the generated position data.

According to this aspect, even when the main scanning reference pattern on the recording medium is partially destroyed by, e.g., contamination, the destroyed portion is

limited to a local area, i.e., a partial image data piece, and the unknown position of the main scanning reference pattern in this local area can be determined by positions represented by feature pixel groups obtained from other normal main scanning reference pattern segments. The positions of the main scanning reference pattern before destruction can be restored with high precision. Therefore, black and white meshes of the mesh pattern can be reliably recognized with reference to the positions of the main scanning reference pattern.

According to one preferred aspect of the present invention, the main scanning reference pattern consists of two continuous lines (guide lines) extending on both sides of a mesh pattern. As a pixel group featuring a piece of this guide line, three, e.g., white, black, and white run lengths (to be referred to as a guide line set hereinafter) which are arrayed in a depth direction of image data (in a direction perpendicular to a main scanning line image) are used. The guide line set is searched on partial image data segments of each of the two guide lines. The position of one of the guide line portions (often both guide line portions), for which search is failed is obtained by interpolation from the normal (search is successful) guide line positions before and after this line portion. Such two guide lines can cope with a change in scanning direction caused when an image sensor is manually scanned. Since the pattern itself is an elongated continuous line and has a feature which is not present in other mesh image elements, it is advantageous for recognition. Furthermore, even when images, such as characters, which are not preferable for decoding, are arranged around an encoded image, the guide lines can serve as boundaries for an encoded image.

The present invention can be applied to a main scanning reference pattern having an arbitrarily appropriate pattern. The following patterns are more preferable since recognition of the main scanning reference pattern can be facilitated: (a) a pattern having a positional feature (e.g., present outside a mesh pattern); (b) a pattern which is unique by itself, and can be easily distinguished from meshes of a mesh pattern.

A data start mark may be formed at a position (before a mesh pattern) scanned prior to a mesh pattern during scanning an encoded image by a line image sensor, and a data end mark may be formed at a position scanned after the mesh pattern. In this case, when these data start and end marks are detected from image data during a decoding operation, more reliable data decoding is assured. For example, when an end portion of a main scanning reference pattern (e.g., the above-mentioned guide line) is destroyed, if an encoded image format is determined such that the data start and end marks are located on the main scanning line or a line group crossing the end portion of the main scanning reference pattern, the destroyed end portion of the main scanning reference pattern can be determined from a line of the main scanning line image from which a piece of the data start or end mark is detected and a line estimated from detection results of undestroyed main scanning reference pattern portions.

The size (depth) of partial image data as a unit area for searching a main scanning reference pattern piece or the size of a piece to be searched may be localized according to a searching result, and may be reduced stepwise until it is almost equal to, e.g., the size of a destroyed piece. For example, when searching is unsuccessful in two continuous partial image data, a piece to be searched is halved to retry searching. In this manner, the position of the main scanning reference pattern can be determined with still higher precision.

According to another aspect of the present invention, image data representing an image is read from a recording

medium which records the image including a mesh pattern and a scanning reference pattern having a synchronous mark synchronous with the mesh pattern. The synchronous mark of the scanning reference pattern is searched from the read image data. During this searching, at least one synchronous mark is detected to determine feature parameters including its position. The position of an adjacent synchronous mark is predicted from the determined feature parameters. The adjacent synchronous mark is actually measured with reference to the predicted position. The position of the actual measurement result is compared with the predicted position. When the comparison result essentially represents a coincidence, the position of the adjacent synchronous mark is determined in accordance with the actual measurement result. However, when the comparison result essentially represents a noncoincidence, the position of the adjacent synchronous mark is determined on the basis of the determined feature parameters. The prediction, actual measurement, comparison, and determination operations are repeated for all the synchronous marks, thereby generating position data of all the synchronous marks. Black and white meshes of the mesh pattern are identified based on the generated position data.

According to this aspect, a portion of the scanning reference pattern, which is destroyed by, e.g., contamination can be detected since the predicted point obtained in the prediction, actual measurement, and comparison operations does not coincide with the actually measured point. In this case, the position of the destroyed synchronous mark is determined based on already determined feature parameters (e.g., an interval or position of the synchronous mark obtained from a mark near the destroyed portion). Therefore, even when the scanning reference pattern is partially destroyed, original synchronous mark positions can be obtained with high precision. Thus, black and white meshes of a mesh pattern can be reliably identified on the basis of the obtained position data.

The scanning reference pattern may be one or both for main scanning and sub-scanning.

When a scanning direction or a scanning speed is varied during scanning of an encoded image by an image sensor, if the position of a synchronous mark is determined on the basis of an actual measurement result, the feature parameters serving as a standard for the next synchronous mark are preferably updated by the actual measurement result.

The meaning of a "position" used when an actually measured position and a predicted position are compared should have a broad sense, and the "position" can include a parameter whose value is changed in association with destruction. For example, an interval between adjacent synchronous clocks is varied when the central position of one synchronous mark is considerably changed by contamination. Thus, a predicted synchronous clock interval does not coincide with an actually measured synchronous clock interval. As for the size of a synchronous mark, a destroyed synchronous mark can be detected by comparison since its size is considerably different from a predicted size of a regular synchronous mark.

When a moving direction (sub-scanning direction) of a line image sensor for scanning an encoded image is varied during scanning, it is preferable that at least two arrays of synchronous marks arranged along an average sub-scanning direction of an encoded image are prepared, and are arranged to be separated in the main scanning direction. When the two arrays of synchronous marks are used, the synchronous marks in the two arrays of synchronous marks

can have a one-to-one correspondence with each other. Therefore, by utilizing this nature, a position of a portion where a synchronous mark in one array is destroyed can be determined with high precision on the basis of an actually measured position of a corresponding synchronous mark in the other array, as well as on the basis of the relationship (position vector) between actually measured positions of adjacent normal synchronous marks in the two arrays of synchronous marks (since an interval between adjacent synchronous marks on each array can be minimized, the change is very small even if the scanning direction is changed at the interval).

According to another aspect of the present invention, in order to record, on a recording medium, an encoded image in which image elements expressing elements of data are arranged at two-dimensionally distributed positions on the recording medium, i.e., an encoded image constituted by a two-dimensional matrix of image elements representing data units, an error checking code is added to a first or original two-dimensional (data) matrix constituting target data to form a second two-dimensional matrix, and an encoded image is printed on the recording medium according to the formed second two-dimensional matrix.

Therefore, since an encoded image on the recording medium includes an image element group representing an error checking code in addition to target data which is significant as original data, the error checking code can be effectively used during data reproduction.

More specifically, in a data reproduction method and apparatus according to this other aspect of the present invention, the above-mentioned encoded image is read from a recording medium as image data, image elements are found from the read image data, data elements expressed by the image elements are obtained to form a two-dimensional data matrix, and an error on the matrix is corrected using an error checking code included in this matrix, thereby reproducing target data correctly.

Furthermore, in a data recording method and apparatus according to still another aspect of the present invention, a means for taking a countermeasure against a burst error is added to the data recording method and apparatus described earlier. The order of elements of a two-dimensional data matrix is changed so that continuous data elements on the two-dimensional data matrix obtained by adding an error checking code to target data are arranged as scattered image elements on a recording medium. An encoded image is recorded according to the two-dimensional data matrix in which the order is changed, so that a two-dimensional data matrix whose order coincides with that of the two-dimensional matrix of image elements on the recording medium is determined as one not before the order change but after the order change.

When the order of the matrix is changed (scrambled or interleaved), even if a contamination as a cause of a burst error is attached onto the recording medium, this portion corresponds to discontinuous data elements on a data matrix before the order change. Therefore, a burst error on image data can be removed during data reproduction.

More specifically, in a data reproduction method and apparatus according to still another aspect of the present invention, image elements are found from read image data to form a two-dimensional data matrix corresponding to a data matrix whose order is changed, as described above, i.e., a data matrix whose order coincides with the order of image elements on the recording medium. Thereafter, the order of the matrix is restored (descrambled or deinterleaved) to an

original order, thereby forming a two-dimensional data matrix corresponding to a data matrix obtained when an error checking code is added to target data. An error on this matrix is corrected based on the error checking code included in the matrix, thereby reproducing the target data.

As a result, even if a burst error is included in a data matrix when the data matrix whose order coincides with the order of elements of an encoded image on the recording medium is obtained, the error appears at separate locations on a descrambled data matrix including the error checking code. Therefore, the error can be easily corrected based on the error checking code.

Furthermore, in a data recording method and apparatus according to a still further aspect of the present invention, when an encoded image constituted by a two-dimensional matrix of image elements which express data elements by black and white meshes is recorded on a recording medium, the values of elements on a first or original two-dimensional data matrix are changed according to a pseudo random number to form a second two-dimensional matrix (randomized data matrix) so that data elements successively having the same value on the first two-dimensional data matrix constituting given data are not recorded as image elements successively having the same density on the recording medium. The encoded image is printed on a recording medium in accordance with the formed second two-dimensional matrix.

Therefore, image elements having the same density hardly continuously appear on the recording medium, and a set of image elements, densities or gray levels of which are irregularly changed is obtained. This irregularity is advantageous for characteristics of an optical image sensor when an encoded image is read by a line image sensor, and can prevent an error caused by a change in photoelectric conversion characteristics of the image sensor caused by long portions having the same density.

In a data reproduction method and apparatus according to a still further aspect of the present invention, which reads such an encoded image, image elements are found from read encoded image data to form a randomized data matrix having data element values according to densities of the image elements, and thereafter, the values of the elements of the matrix are changed according to a predetermined pseudo random number (derandomized) to obtain a data matrix before randomization, thereby reproducing data.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Other objects and features of the present invention will be understood for those skilled in the art from the description of the preferred embodiments of the present invention taken in conjunction with the accompanying drawings, in which:

FIG. 1 is a block diagram of a data reading apparatus according to the first embodiment of the present invention;

FIG. 2 exemplifies an encoded data on a recording medium as an object to be read;

FIG. 3 exemplifies an image obtained by scanning the encoded image shown in FIG. 2 by a manual image scanner;

FIGS. 4A and 4B show a flow chart for storing image data obtained by scanning an encoded image, and recognizing a guideline as a main scanning reference pattern from the stored image data to perform main-scanning decoding;

FIGS. 5A and 5B show a flow chart for recognizing a synchronous mark array as a sub-scanning reference pattern from the main-scanning decoded image data to perform sub-scanning decoding;

FIG. 6 shows a memory map of an image RAM;  
FIGS. 7(a) and 7(b) are views for explaining a guide line set featuring a piece of a guide line;

FIG. 8 shows an encoded image around which characters are present, and a plurality of search blocks obtained by dividing the encoded image;

FIG. 9 exemplifies an image including a contaminated guide line and a synchronous mark array;

FIG. 10 shows main-scanning decoded image data corresponding to the image shown in FIG. 9;

FIG. 11 is a block diagram of a data reading apparatus according to the second embodiment of the present invention;

FIG. 12 shows an encoded image having no data end mark;

FIG. 13 is a schematic view of an encoded image with three guide lines;

FIG. 14 is a schematic view of an encoded image with one guide line;

FIG. 15 is a schematic view of an encoded image with one synchronous mark array;

FIG. 16 is a schematic view of an encoded image with synchronous mark arrays in which elongated black marks are arrayed as clocks;

FIG. 17 is a schematic view of an encoded image with synchronous mark arrays having different cycles;

FIG. 18 is a view for explaining processing for identifying black and white meshes by a middle point method from the center of four clocks of the synchronous mark arrays;

FIG. 19 shows an arrangement of image data on an image RAM including middle points C1 and C2 shown in FIG. 18;

FIG. 20 shows the relationship between four positions of a sampling reference pattern and a mesh position located inside a rectangle defined by the four positions on read image data;

FIG. 21 shows the relationship between four positions of a sampling reference pattern and a mesh position located inside a rectangle defined by the four positions on a recording medium in correspondence with FIG. 20;

FIGS. 22(a) - 22(f) views for explaining processing for coordinate-converting the central coordinates of meshes on a recording medium into the central position of meshes on image data;

FIG. 23 is a block diagram showing a data recording apparatus according to the third embodiment of the present invention;

FIG. 24 is a flow chart showing the overall operation of the data recording apparatus shown in FIG. 23;

FIG. 25 shows a bit matrix of target data;

FIG. 26 shows a symbol matrix of the target data;

FIG. 27 shows a matrix in which an error checking code is added to the target data;

FIGS. 28A and 28B show a matrix after a first scrambling operation of a data matrix with the error checking code;

FIGS. 29A and 29B show a matrix after a second scrambling operation;

FIGS. 30A and 30B show a matrix after a third scrambling operation;

FIG. 31 shows a scrambled bit matrix;

FIG. 32 shows an arrangement of a pseudo random number generator for removing a DC component from a scrambled bit matrix;

FIG. 33 shows a bit matrix from which a DC component is removed, i.e., a final bit matrix;

FIG. 34 is a flow chart showing the overall operation of a data reproduction apparatus;

FIG. 35 shows a symbol matrix of data with an error checking code after-symbol descrambling during data reproduction together with P and Q vectors processed in error correction processing;

FIG. 36 is an overall flow chart of error correction processing;

FIG. 37 is a detailed flow chart of steps 24-2 and 24-5 of the processing shown in FIG. 36;

FIG. 38 is a detailed flow chart of erasure processing shown in FIG. 37; and

FIGS. 39(a) - 39(g) show encoded images with errors which data can be reproduced by the data reproduction apparatus of the third embodiment.

#### DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

Preferred embodiments of the present invention will be described hereinafter with reference to the accompanying drawings.

##### First Embodiment

FIG. 1 shows the overall arrangement of a data reading apparatus 10 according to the first embodiment. The object of this apparatus is to read a mesh-pattern image 20 (FIG. 2) recorded on a recording medium such as a paper sheet using an image sensor 11 and to decode data encoded in the image of FIG. 2.

The image sensor 11 comprises a manual line image sensor including a CCD sensor element array. The image sensor 11 is manually moved in a sub-scanning direction (lateral direction) of the recording medium to scan the encoded image 20 on the recording medium, thereby generating corresponding image data. More specifically, the sensor elements of the sensor array photoelectrically convert input light, i.e., input reflected light according to a black-and-white pattern of pixels on the recording medium facing the sensor elements into analog signals. Analog image outputs of the sensor elements are sequentially and cyclically derived through an analog multiplexer or the like, so that a line image output for all the sensor elements of the sensor array can be obtained for each predetermined period (main scanning period) determining a time for one main scanning line. Thereafter, the line image output is binarized by a binary circuit (e.g., a black pixel="1" and a white pixel="0"), and the binary serial image data is supplied to a controller 12 together with a timing control signal. The controller 12 controls the image sensor 11 and its peripheral circuits, and performs conversion of the input serial image data into parallel data.

A CPU 13 operates according to a program stored in a ROM 14. During scanning of an image by the image sensor 11, the CPU 13 writes parallel data in an image RAM 15 every time the serial image data is parallel-converted by the controller 12. Upon completion of the image scanning, the CPU 13 starts a decoding operation of the image data stored in the image RAM 15, and records the decoding result in a RAM 16. The decoding operation can be roughly classified into processing for recognizing a sampling reference pattern of data from the image data, and processing for identifying black and white meshes of a mesh pattern as a data body in accordance with the recognition result. According to this embodiment, even if the sampling reference pattern is partially destroyed by, e.g., contamination, the respective posi-

tions of the sampling reference pattern on the image data can be determined with high precision, and an accurate decoding result of the mesh pattern can be obtained. This will be described in detail later.

FIG. 2 shows the encoded image recorded on the recording medium as an object to be read by the data reading apparatus 10. The encoded image 20 shown in FIG. 2 includes a mesh pattern 22 as a data body. The mesh pattern 22 consists of a large number of meshes arranged in a matrix. Each mesh is selectively set in black or white to express (encode) a bit as a unit of data. For example, a black mesh expresses a bit "1", and a white mesh expresses a bit "0". In FIG. 2, the mesh pattern 22 has 48×72 bit information. Furthermore, the encoded image 20 has patterns or marks for indicating main scanning and sub-scanning data sampling references for the mesh pattern 22 as a data body. In FIG. 2, the main scanning and sub-scanning sampling reference marks are separately formed. The main scanning sampling reference is defined by two guide lines 21 extending in the sub-scanning direction along the upper and lower sides of the mesh pattern. The sub-scanning sampling reference is defined by two synchronous mark arrays 25 each of which extends along the inside of the corresponding guide line 21 and has repetitive patterns of alternate black and white. An interval or pitch between adjacent black and white meshes in each synchronous mark array 25 is synchronous with, i.e., coincides with a sub-scanning interval of meshes in the mesh pattern 22. Thus, each mesh of the synchronous mark array 25 will be called a "clock" hereinafter.

The reason why the two guide lines 21 are formed as the main scanning reference patterns along the moving direction of the image sensor 11, i.e., the sub-scanning direction of the image 20 to be separated at a predetermined distance, is to cope with a case wherein the moving direction of the image sensor 11 is changed during manual, scanning of an image. More specifically, image data from the image sensor 11 is stored in the image RAM 15 to be identified from each other in units of main scanning lines. A direction of a line image represented by each main scanning line image data cannot be clearly determined. However, when the positions of the two guide lines 21 on the main scanning line image data are detected, a position (central position of main scanning data sampling) on each main scanning line image data corresponding to the vertical central position (or may be a vertical mesh range) of meshes in the mesh pattern 22 can be determined based on the detection results in association with the positional relationship of an original image, i.e., the predetermined positional relationship formed between the two guide lines 21 and the mesh pattern 22 in the encoded image on the recording medium as an original. For example, in the image format shown in FIG. 2, the two guide lines 21 are parallel to each other, and are also parallel to the mesh pattern 22. Since a straight line connecting one point in the upper guide line 21 and one point in the lower guide line 21 passes a plurality of meshes arranged at equal intervals in the mesh pattern 22, line image data expressing an image of such a straight line should also have the same feature. Therefore, points of the two guide lines 21 are detected from the line image data, and an interval between the two points is equally divided to obtain the main scanning central position of meshes in the line image.

The synchronous mark arrays 25 as the sub-scanning reference patterns are utilized to determine the positions of meshes in the sub-scanning direction in the read image data. More specifically, clocks of the upper and lower synchronous mark arrays 25 are detected from the image data, and corresponding upper and lower clocks are checked. In the

encoded image format shown in FIG. 2, a straight line connecting the centers of these corresponding clocks passes the center in the lateral direction (sub-scanning direction) of meshes in a certain column of the mesh pattern 22. Therefore, the positions of clocks in the pair of synchronous mark arrays 25 are detected from the image data, thus determining the sub-scanning central position or a sub-scanning mesh range of meshes in the mesh pattern 22 on the image data.

In the encoded image 20 shown in FIG. 2, a data start mark 23 consisting of a checkerboard mesh pattern is formed in an area scanned prior to the mesh pattern 22. A data end mark 24 in which laterally elongated black and white stripes alternately appear along the main scanning direction is formed in an area scanned after scanning of the mesh pattern 22. Thus, since the data start and end marks 23 and 24 comprise periodic or regular patterns, they can be easily detected on the image data. Since the data start and end marks 23 and 24 respectively comprise patterns which can be easily visually distinguished from each other, a user can easily determine the scanning direction (moving direction) of the image sensor 11 with respect to the encoded image 20. Even when an end portion of the guide line 21 as the main scanning reference is destroyed by, e.g., contamination, start or end of data can be detected upon detection of the data start or end mark 23 or 24, and the destroyed position of the guide line 21 can be predicted or evaluated.

In FIG. 2,  $W1=W2=W4=W5$ . Where  $W1$  represents a distance between a start end of the guide lines 21 and a start end of the data start mark 23,  $W2$  a width of the mark 23,  $W4$  a width of the data end mark 24, and  $W5$  a distance between an end terminal of the mark 24 and ends of the guide lines 21.  $W3$  designates the length of the data body.

FIG. 3 shows an image obtained when the encoded image shown in FIG. 2 is manually scanned by the image sensor 11 relatively carefully. Even when the sampling reference pattern is partially omitted, i.e., when the guide lines 21 or the synchronous mark arrays 25 are slightly destroyed, the apparatus of this embodiment can cope with this and can provide a correct decoding result as long as a read image is distorted to a degree as shown in FIG. 3.

Image data decoding processing according to this embodiment will be described below.

FIGS. 4A and 4B show a flow chart of processing for storing image data obtained by scanning the encoded image 20, and recognizing the guide lines 21 as main scanning reference patterns from the stored image data, thereby main-scanning sampling image data. FIGS. 5A and 5B show a flow chart of processing for recognizing the synchronous mark arrays 25 as the sub-scanning reference patterns from the main-scanning sampled image data, and sub-scanning sampling the image data. This embodiment has a characteristic feature in recognition of the guide lines 21 and the synchronous mark arrays 25. Even when the guide lines 21 or the synchronous mark arrays 25 are slightly destroyed, the destroyed position can be determined with high precision.

In steps 4-1 to 4-5 in FIG. 4A, the encoded image 20 on the recording medium is read by the image sensor 11, and the read image data is written in the image RAM 15. A scanning end condition (memory full) in step 4-3 is merely an example, and scanning end may be determined in response to various other appropriate events. The image RAM 15 is assumed to comprise a byte memory as in step 4-4. FIG. 6 shows a memory map of the image RAM 15. A one-line image (line image) is written in one of 1 horizontal lines (1 successive addresses of the image RAM 15) in FIG. 6.

The image decoding operation starts from step 4-6 in FIG. 4A. In step 4-6, a guide line set (pixel group featuring the guide line as the main scanning reference) is searched from the entire image data stored in the format shown in FIG. 6. As shown in FIG. 2, the guide line 21 is a black continuous line, and has a feature which is not present in other elements of the encoded image 20. Therefore, for example, the guide line set can be defined by three, i.e., white, black, and white parallel run lengths 73, 74, and 75 having an appropriate interval, as shown in FIGS. 7(a)-7(b). An initial value of an interval or width 76 of the run lengths necessary for finding the guide line set may employ a fixed standard limit value or may be determined by measuring the widths of black or white meshes which most frequently appear in a line image. The guide line set defined by the three appropriate run lengths 73, 74, and 75 can be searched as follows. That is, an image is traced from the positions of white, black, and white meshes having an appropriate interval on an appropriate line image on the image RAM 15 in a depth direction (vertical direction in FIG. 6) perpendicular to the line image to detect the numbers of successive white, black, and white meshes (run lengths), and the detection results are compared with a condition of the guide line set. This processing is repeated. When searching is unsuccessful, an error occurs since decoding is impossible (4-7). However, when searching is successful, the width of the guide line and the standard value of an inclination of the scanning direction are determined based on the information of the guide line set. In step 4-8, margins taking a variation from the standard value of the inclination of the scanning direction into consideration are added to the right and left (upper and lower in FIG. 2, but "right and left" will be used according to FIG. 6) guide line set positions, thereby determining right and left searching areas (searching widths 81 in FIG. 8) of image data in the following processing. It is necessary to define the searching widths 81 when images such as characters and the like (which may cause noise during decoding) are present around the encoded image 20, as shown in FIG. 8. However, this processing is not necessary when a blank space is formed around the encoded image 20.

In steps 4-9 to 4-11, processing for defining a searching area of image data in a depth direction is executed. For this purpose, the data start and end marks 23 and 24 are detected to obtain left and right guide line positions where these marks 23 and 24 are detected. The data start and end marks 23 and 24 can be detected by detecting a line image including periodic black and white patterns. For example, in the encoded image shown in FIG. 2, the following simple condition for detecting these marks can be sufficiently used. That is, since each of the start and end marks 23 and 24 consists of 24 pairs of black and white repetitive patterns, when about 20 pairs of black and white patterns having the same cycle are found in consideration of a margin, it can be determined that these marks are detected. The data start mark 23 is preferably searched from an upper line (first line) of the image data, and the data end mark 24 is preferably searched from a lower line (last line) of the image data. When the start or end mark 23 or 24 cannot be detected, error processing is executed since it can be considered that the error is caused by an erroneous operation such as partial scanning of the encoded image 20 (4-10 or 4-12).

(18)  
In steps 4-13 to 4-18, image data in which the searching width 81 and the searching depth (from a start line 82 to an end line 83 in FIG. 8) are defined by the above-mentioned processing is divided into a plurality of partial image segments in the depth direction, which are called searching blocks, and whether or not the guide line in each segment is

destroyed is checked. In FIG. 8, the image data is divided into eight searching blocks 84. The size (depth) of the searching block is determined in step 4-13, and the first searching block starting from the line next to the start line 82 is selected. The depth of the searching block can be determined in consideration of the guide line set information obtained in step 4-6, a processing time, precision, and the like, and this value determines a necessary length of the white, black, and white run lengths associated with the guide line set searched in step 4-14. More specifically, when run lengths having an appropriate interval and having a length equal to or larger than the depth of the searching block are detected as the white, black, and white run lengths in the depth direction, it can be determined that the guide line sets are present, and data representing their positions (a rectangular area surrounded by the guide line sets) is stored. When searching is unsuccessful, since it indicates that the guide lines of the searching block are damaged, a flag according to an unsuccessful state indicating for which one of the right and left guide lines searching is unsuccessful or that searching is unsuccessful for both the guide lines is set so that guide line positions can be interpolated later (4-15, 4-16). For example, in FIG. 8, since a contamination 85 is attached to the right guide line portions of the fourth and fifth searching blocks from above, the right guide line portions in these searching blocks are detected as errors.

After the state of the guide lines is checked for one searching block, the next searching block is selected (4-17), and checking is repeated until the last searching block at the end line 83 (4-18) is checked.

Although not executed in the flow shown in FIG. 4A, the size of each searching block may be variably localized according to the searching result. For example, if a guide line error is detected in a certain searching block in step 4-15, the depth of the searching block is halved, so that guide line searching is executed for the two half-depth searching blocks again or the loop including steps 4-13 to 4-18 may be restarted using half-depth searching blocks.

From step 4-19 to the last step 4-31 in FIG. 4B, the right and left guide line positions in each main scanning line image are determined for each searching block, and a main scanning sampling position of each line image (the entire path of the sampling positions is denoted by reference numeral 31 in FIG. 3) is obtained by an equal dividing method. Image data bits (pixels or meshes) at the sampling positions are extracted from the line images to form main-scanning decoded arrays 100 (FIG. 10). The guide line positions (central positions in the guide line widths) on each main scanning line image are determined by actual measurement when no failure flag indicating an error or destruction is set for the guide line portions of a searching block to which the main scanning line image belongs. However, when the failure flag is set, the guide line positions are determined by interpolation based on other guide line positions actually measured from normal guide line portions. For example, as shown in FIG. 4B, a list of failure flags are looked up, and guide line positions in a main scanning line image of interest can be obtained by linear interpolation from Guide line positions in main scanning line images at the ends of previous and following searching blocks.

According to the flow of FIG. 4B, NO is determined in step 4-20 when both the right and left guide line portions of a current searching block are correct (it can be determined since the corresponding failure flags are reset). In this case, the central positions of the right and left guide line portions of each main scanning line image in the current searching block are actually measured in step 4-29, and an equally

divided position according to the encoded image format between the two central positions is then obtained as a main scanning sampling point. Thereafter, an image bit at each sampling point is extracted. If it is determined in step 4-20 that the failure flag of at least one of the right and left guide line portions is set, the flow advances to step 4-21 and subsequent processing. In order to interpolate the guide line position for which the failure flag is set, the guide line position immediately preceding the current searching block, i.e., the guide line position in the last main scanning line image in the previous searching block (or the guide line position of the start line obtained in step 4-9 if the current searching block is the first searching block) is determined as an interpolation start point (4-21). A searching block including normal guide line portions is searched from the following searching blocks (4-22, 4-26), and the guide line position in the uppermost main scanning line image in the normal searching block is detected as an interpolation candidate (4-27). The guide line positions on the main scanning line images in the searching block of interest present between the interpolation start and end points are determined by interpolation, and main-scanning sampling is then executed (4-28). Although not shown in the flow, when one guide line portion in a certain searching block is normal, its position is directly measured. When a guide line portion near the end line is destroyed, the guide line position evaluated in step 4-11 is used as an interpolation end point (4-23, 4-24).

A guide line section including an error is interpolated on the basis of the positions actually measured in other normal guide line sections, thus executing high-precision main-scanning sampling. As a result, the arrays of the main-scanning decoded image data 100 are completed, as shown in FIG. 10. Columns on two sides of the main-scanning decoded image data 100 are linear image mesh arrays extending along the main scanning center path 31 (FIG. 3) of the synchronous mark arrays 25 as the sub-scanning scanning reference patterns. In the flow chart of FIGS. 5A and 5B, the mesh arrays of the right and left synchronous mark arrays 25 are checked to determine the sub-scanning central positions of the clocks in the synchronous mark arrays 25, and the mesh pattern 22 is then subjected to sub-scanning data sampling to identify black and white levels of meshes.

In step 5-1 in FIG. 5A, the length in the depth direction of a clock serving as a standard value (comparison reference value) of the start clock in step 5-3 is determined on the basis of the main scanning width, i.e., an interval between the right and left guide lines 21 (obtained in FIGS. 4A and 4B). In this embodiment, the image sensor 11 is assumed to be manually scanned on the encoded image 20, and a clock length may be considerably varied. Therefore, a relatively large margin must be added to the standard value. Note that in place of calculating the standard value from the main scanning width, the mesh arrays of the synchronous mark arrays 25 on the main-scanning decoded image data 100 are checked to obtain an average black or white run length, and the obtained run length may be used as the standard value. In step 5-2, the first (black) clocks of the right and left synchronous mark arrays 25 on two sides are detected from the main-scanning decoded image data, and their central points, lengths, an inclination between the right and left first clocks with respect to the main-scanning direction of the image data 100 (horizontal direction in FIG. 10), and the like are actually measured. In step 5-3, it is checked if the clock length of the measurement results falls within the range of the standard value obtained in step 5-1. When the clock length does not fall within the range of the standard value at

this time, a read error occurs. When the clock length falls within the range of the standard value, values of meshes on a straight line connecting the central positions of the right and left start clocks are extracted from the main-scanning decoded image data 100 to obtain data representing black and white levels of meshes of the first column in the mesh pattern 22 in order to perform sub-scanning sampling. In addition, the feature parameters (the central positions, the lengths, the inclination, and the like) are set as the standard values for the next clocks.

In step 5-5, the length of the previous clock is added to the central position of the previous clock to predict the central position of the next clock. In this embodiment, this prediction can be made since the synchronous mark arrays of the encoded image 20 (FIG. 2) comprise repetitive patterns of black and white clocks having the same length, and hence, the clock length is equal to the clock interval. In step 5-6, searching is vertically performed from this predicted point along the mesh array of the synchronous mark array on the main-scanning decoded image data until a mesh value different from that of the predicted point is detected, thereby actually measuring the next clock. For example, if the predicted point is "1" representing a black pixel, vertically successive "1"s having the predicted point as the center are determined as the next clocks. The length, the center, an inclination between the right and left clocks, and the like of the next clocks are obtained as actual measurement results (5-7). When prediction and actual measurement are executed in this manner, if a clock includes no error, the predicted central point and the actually measured central point should fall within a predetermined range. However, if the clock is destroyed, its clock length and inclination must be considerably changed from the standard values (the length of the previous clock and the inclination between the previous right and left clocks), as shown in FIGS. 9 and 10. In FIG. 9, since a contamination 63 paints a black clock 25a of the right synchronous mark array 25, a black mesh array indicating a piece of the contamination 63 surrounded by a meshed line C in FIG. 10 is formed in the mesh array (pixel array on the scanning path 31 shown in FIG. 9) of the synchronous mark array 25 on the main-scanning decoded image data 100 shown in FIG. 10. Therefore, the length of the next clock is observed to be longer by the contamination than that of the previous clock (white clock and its actual measurement center is indicated by P1 in FIG. 10) in the mesh array of the right synchronous mark array 25. A predicted center P2 of the next clock is largely deviated from an actually measured center P3. Since a corresponding portion of the right synchronous mark array 25 has no contamination, a difference between a center P2 predicted from the previous clock (its center is indicated by P1) and an actually measured value P3 of the next clock is very small if there is any. Therefore, the feature parameters predicted for the next clock are compared with actually measured feature parameters, and a difference therebetween is checked to detect a clock error caused by, e.g., contamination. In the flow chart of FIG. 5B, a difference between the predicted central point of the next clock and the actually measured central point of the next clock for each of the right and left synchronous marks is obtained (5-8), and it is checked if the difference falls within an allowable range (5-9, 5-10, 5-13), thereby discriminating whether or not the next clock is proper. Step 5-11 is executed when neither the right nor left clocks are proper. In this situation, since only the previous state of the synchronous mark array 25 is reliable, the right and left predicted points obtained in step 5-5 are determined as the central points of the next clocks. In step 5-16, a mesh

array on a straight line connecting the central points is sampled from the main-scanning decoded image data 100, and the sampled data array is determined as data indicating black and white levels of associated meshes. Steps 5-12 and 5-14 are executed when one of the right and left next clocks is proper, and the other is not proper. In this case, the central point of the improper clock may employ a predicted point from the previous clock. However, in order to improve precision, the central point of the improper clock is obtained from the actually measured central point of the proper clock on the basis of the inclination. For example, the inclination of right and left proper white clocks observed in an upper portion of FIGS. 9 and 10 can be evaluated by a vertical difference of 2 (meshes) between the central points P1 of right and left clocks. Of black clocks located immediately below these two white clocks, the right black clock is improper, and the left black clock is proper. The central point of the left black clock is actually measured as P2 in FIG. 10. Therefore, a position two meshes above an intersection between a horizontal line passing through the actually measured point and a straight line as the mesh array of the right synchronous mark array 25 (in this case, this position fortunately coincides with the position P2 predicted from the center P1 of the right previous clock) is determined as the central point of the right improper clock. In step 5-15, only the clock length of the proper clock is updated as the standard value of a clock interval. In step 5-16, a mesh array between the centers of the right and left clocks is sampled. When both the clocks are proper, the actually measured clock centers are settled, and a mesh array therebetween is extracted. In this case, all the standard parameters are updated by feature parameters of the actually measured clocks in step 5-15.

In step 5-17, end of decoding is checked. In this step, the number of data determined according to the format of the encoded image is compared with the number of execution times of decoding step 5-16. When the number of execution times of step 5-16 is equal to the number of data determined by the format, and the mesh array of the following synchronous mark array 25 does not include a clock up to the end line (5-18), i.e., when the number of clocks determined by actual measurement or interpolation from the mesh array of the synchronous mark array 25 on the main-scanning decoded image data 100 is equal to the number defined by the format, it is determined that proper processing operations are executed, thus completing sampling processing. When erroneous clock interpolation or the like is executed during sampling processing, the end line of the image data arrays 100 is reached before the number of data defined by the format is obtained (NO in step 5-19), or another clock is found after the number of data defined by the format is obtained (NO in step 5-18). Thus, an error can be detected.

As described above, according to the first embodiment, sampling reference patterns (the guide lines 21 and the synchronous mark arrays 25) included in the encoded image 20 are improved in consideration of reading of the encoded image 20 under a considerably strict environmental condition, e.g., a change in scanning speed or direction during scanning which often occurs since the encoded image 20 is manually scanned by the image sensor 11, characters and the like which cause noise and are recorded around the encoded image 20, and the like. In addition, a countermeasure against a case wherein the sampling reference patterns are partially destroyed is taken in recognition of the sampling reference patterns in the read image data. That is, for the guide lines 21 as the main scanning reference patterns, measurement of the positions of the guide lines 21 in a

portion other than the destroyed portion is allowed by a segment method, and the guide line position of the destroyed portion is interpolated based on the obtained position information. For the synchronous mark arrays 25 as the sub-scanning reference patterns, the positions of the clocks are obtained by prediction based on feature parameters of an adjacent clock, actual measurement from the predicted point, comparison between the actual measurement result and the prediction result, and predetermination of a position whose comparison result indicates a clock error. Therefore, the positions of the sampling reference patterns can be determined from image data with higher precision even in a relatively strict use environment, e.g., the presence of misleading characters, a variation in scanning direction, contamination, and the like, and black and white levels of meshes of the mesh pattern can be identified with high precision on the basis of the determination result.

#### Second Embodiment

FIG. 11 shows the overall arrangement of a data reading apparatus 110 according to the second embodiment of the present invention. In this embodiment, an image sensor 111 mechanically scans a recording medium 118. As shown in FIG. 11, a controller 112 drives a stepping motor 117. The stepping motor 117 moves the image sensor 111 via pulleys 117a, 117b and a belt 117c along a predetermined path, e.g., a rail. With this arrangement, since variations in moving speed and direction of the image sensor 111 are considerably reduced as compared to a manual operation, a distortion of read image data can be remarkably reduced (such a merit is attained even when the motor 117 is not controlled by a special servo mechanism or the like). Therefore, by utilizing this nature, some of constituting elements and processing operations described in the first embodiment can be facilitated or omitted. For example, even if an encoded image 120 having no data end mark shown in FIG. 12 is used, the number of lines below a data start mark 23 until data ends can be roughly predicted based on a precise scanning speed of the image sensor 111. Even if the size of each searching block is increased, the position of a destroyed guide line portion can be interpolated with high precision. In recognition processing of synchronous mark arrays 25 as sub-scanning reference patterns, not local parameters, e.g., feature parameters of an immediately preceding clock but global parameters averaged from the entire image can be used as comparison standard values. Thus, the position of the next clock can be predicted and evaluated based on the global parameters with high precision.

#### Another Arrangement

The first and second embodiments have been described above. Various changes and modifications of the present invention may be made within the spirit and scope of the invention.

The data start and end marks 23 and 24 described in the first embodiment are effectively used to predict their corresponding portions (executed in steps 4-9 and 4-11 in FIG. 4A) when the corresponding portions of the marks are destroyed. However, the normal functions of the marks 23 and 24, i.e., the functions of indicating start and end of data can be realized without using these marks. For example, if the data start and end marks 23 and 24 are omitted from FIG. 2, successive white pixels appear between black pieces of the two guide lines 21 in a main scanning line image obtained by scanning these portions (white portions), thus detecting start and end of data.

In the above embodiments, the two guide lines 21 are arranged outside the mesh pattern 22 as the main scanning reference patterns in the encoded image 20 or 120. However,

various other appropriate main scanning reference patterns may be used. For example, as shown in FIG. 13, when three (or more) parallel guide lines PR are used as main scanning reference patterns for a mesh pattern MP, when the guide lines PR is partially destroyed, the guide line position can be more strongly restored. For example, when one of the three guide lines PR is partially destroyed, the destroyed guide line portion can be accurately determined based on the corresponding positions of the remaining two guide lines. As exemplified by the central guide line PR in FIG. 13, a main scanning reference pattern may be arranged in the mesh pattern 22. When the scanning direction of the image sensor is stable, the main scanning reference pattern may be constituted by one guide line PR extending along the main scanning direction of the mesh pattern MP, as shown in FIG. 14. In this case, the scanning direction (moving direction) of the image sensor can be determined by checking a main scanning line image passing through the guide line PR. By checking the main scanning line image, the end position of the guide line PR can be determined. Thus, the central position of each mesh in the main scanning direction on the read image data can be determined on the basis of the above determination result and the positional relationship between the end points of the guide line PR determined as an image format and meshes of the mesh pattern MP, thus executing main scanning sampling. Furthermore, it is advantageous that the main scanning reference pattern comprises a continuous line in black or in a color different from surrounding colors, but may comprise a discontinuous mark.

(34)

In the encoded image 20 or 120 of the above embodiments, the two synchronous mark arrays 25 in which black and white meshes (clocks) alternately appear on the two sides of the mesh pattern 22 are used as the sub-scanning reference patterns. However, various other appropriate patterns synchronous with sub-scanning mesh arrays (which can have a positional correspondence) may be used as sub-scanning references. For example, when the scanning direction of the image sensor is stable, one synchronous mark array SR extending along the sub-scanning direction of the mesh pattern MP can be sufficiently used, as shown in FIG. 15. In order to restore partially destroyed clocks with higher precision, three or more synchronous mark arrays may be arranged to be separated at given intervals in the main scanning direction. The length of each clock in a synchronous mark can be arbitrarily determined. For example, the sub-scanning reference patterns may be formed, as shown in FIG. 16. That is, synchronous marks SR in each of which clocks forming an elongated black piece are aligned may be arranged above and below the mesh pattern MP to have the relationship that a line connecting corresponding clocks passes through a boundary between adjacent meshes of the mesh pattern MP. Alternatively, as shown in FIG. 17, in addition to a first clock mark array SR1, a second clock mark array SR2 in which each clock mark is arranged every predetermined number (8, in this case) of clock marks of the first clock mark array SR1 may be provided. With these clock arrays, a sub-scanning reference for the mesh pattern MP may be formed. In this case, the number of clocks can be easily managed in an image decoding operation, and partially omitted clock marks can be more reliably restored. For example, a clock mark counter for the first clock mark array SR1 is initialized at a position where a coincidence between clocks of the first and second clock mark arrays SR1 and SR2 is found on the read image data. Thereafter, every time a clock mark on the first clock mark array SR1 is detected or is determined by restoration according to this embodiment, the clock mark

counter is incremented. When a coincidence between clocks of the first and second clock mark arrays SR1 and SR2 is detected again, the clock mark counter is checked. In the format shown in FIG. 17, if the content of the counter is equal to  $8n$  ( $n$  is a natural number), it can be determined that the first clock mark array is appropriately recognized. If  $n$  is 2 or more, it can be evaluated that the second clock mark array SR2 is partially omitted. If the content of the counter is not equal to  $8n$ , it can be determined that the first clock marks are erroneously recognized, and an error can be detected. As a result, when the first clock mark array SR1 is partially destroyed and clocks are restored through prediction of the next clock position, actual measurement, comparison, and reevaluation processing operations (see FIGS. 5A and 5B) described in the above embodiments, an error occurring in the restoration process can be restricted to a narrow local area as a clock interval of the second clock mark array SR2. Furthermore, clock positions of the restricted portion can be obtained by equally dividing an interval between the two clocks of the second clock mark array SR2. Therefore, black and white levels of meshes of the entire mesh pattern MP can be identified while recognizing such a local error (in the above embodiment, an error is detected in the whole system as shown in steps 5-17 to 5-19 in FIG. 5B, and decoding is disabled). In a practical application, a sufficient correction opportunity can be given to black-and-white identification data at a position where the above-mentioned local error is found in error correction processing using these checking codes.

The above embodiments are made under an assumption that images such as characters which are not suitable for decoding are present around an encoded image of course, such an assumption need not be made. In this case, recognition of sampling reference patterns can be considerably facilitated, and conditions of reference patterns (conditions of arrangements, geometrical shapes, and the like) for distinguishing reference patterns from images such as characters can be rendered less strict.

In the above embodiments, the main scanning sampling reference patterns (guide lines 21) and the sub-scanning sampling reference patterns (synchronous mark arrays 25) are realized by separate constituting elements on an image. However, a sampling reference pattern which can be commonly used as these patterns may be used.

(37)

Referring back to FIG. 16, for example, the synchronous mark array SR may be used as a common main scanning/sub-scanning reference mark. In FIG. 16, the synchronous mark array SR is present outside the mesh pattern MP as a data body. Therefore, if there is no unnecessary image excluding the mesh pattern MP and the synchronous mark array SR, it is not difficult to detect the synchronous mark array SR from the read image data. For example, paying attention to each main scanning line image of the read image data, black pixels appearing at two ends of each main scanning line image suggest a possibility of the presence of pieces of the synchronous mark array SR. In manual scanning, each main scanning line image is normally read not in a perfectly vertical direction but in a slightly inclined direction with some variations. Therefore, if several (two or three) main scanning line images having black pixel groups near their two end portions, successively appear, they indicate a high possibility of the clock marks of the synchronous mark array SR. Of these black pixel groups, black pixels groups which are distributed on image data to have a certain period can be surely determined as clock marks. Thus, paying attention to black pixel groups for several lines, if their centers are actually measured, the measured centers can

be determined as those of clock marks. A destroyed clock mark can be detected as follows as in the corresponding processing in FIGS. 5A and 5B. That is, the position of the destroyed clock mark is predicted from an immediately preceding clock mark and its interval or pitch (a distance between the centers of the immediately preceding clock mark and a clock mark preceding the immediately preceding clock mark), and is actually measured using the predicted position as the center. The predicted value is compared with the actual measurement value to detect if a difference therebetween is large. If the difference is large, the predicted point is determined as the center of the current clock mark, or a position determined according to the center of the center of the current clock mark in another array on the basis of an inclination between preceding clock marks is determined as the center of the current clock mark. When the central position of each clock mark is determined in this manner, corresponding clock marks are selected from the upper and lower synchronous mark arrays SR, and their centers are connected by a straight line. The straight line is equally divided according to the number of image meshes, and image bits present at equally divided positions are read to identify black and white levels at the centers of meshes (the last description assumes a case wherein each clock of the synchronous mark array SR in FIG. 16 is shifted by half a mesh in the horizontal direction to coincide with the center of each mesh of the mesh pattern MP).

As shown in FIG. 16, when the longitudinal direction of each clock coincides with the boundary between adjacent meshes of the mesh pattern MP, centers  $U_i$ ,  $U_{i+1}$ ,  $B_j$ , and  $B_{j+1}$  of  $i$ th and  $(i+1)$ th clock marks are selected from the upper and lower synchronous mark arrays SR, and each central position of a target vertical mesh array can be calculated from these position data. This will be described below with reference to FIGS. 18 and 19. In FIG. 18, the centers  $U_i$  and  $U_{i+1}$  (positions on the read image data) of the  $i$ th and  $(i+1)$ th clock marks in the upper synchronous mark array SR are connected, and their middle point C1 is calculated. Similarly, the centers  $B_j$  and  $B_{j+1}$  of the  $i$ th and  $(i+1)$ th clock marks in the lower synchronous mark array SR are connected to calculate their middle point C2. FIG. 19 shows a portion including the middle points C1 and C2 of the read image data. Each horizontal line in FIG. 19 corresponds to one main scanning line image. One cell in FIG. 19 indicates a 1-bit (1-pixel) memory cell. The middle points C1 and C2 are connected by a straight line, and the line is equally divided according to an image format, thereby determining the central position of meshes in a target mesh array on the image data memory, i.e., a storage position (indicated by a mark "x" in FIG. 19). With the above-mentioned middle point method, since a data sampling point is obtained from two pairs of two adjacent clock centers, errors included in positions  $U_i$ ,  $U_{i+1}$ ,  $B_j$ , and  $B_{j+1}$  as synchronous mark array SR measurement points (or interpolation points) can be absorbed and decreased by averaging.

Points indicated by the marks "x" in FIG. 19 need only be data-sampled to identify black and white levels of meshes. Surrounding pixel values (e.g., upper, lower, right, and left pixel values) of the point indicated by the mark "x" are checked, and black and white levels of meshes may be determined based on the checking results, if so desired. Alternatively, when surrounding pixel values contradict with the pixel value of the point indicated by the mark "x" obtained based on the recognition result of the synchronous mark array SR, data indicating this is stored, and can be effectively used in the error correction processing later using the checking codes, thus enhancing error correction capa-

bility. For example, when an error is detected when black-and-white identification data of a certain block is checked according to a checking word, if the block includes a position where a contradiction is found, this position is predicted as a position where the error occurs, thus allowing so-called erasure correction (error correction when an error position is known).

If a sampling reference pattern and a mesh pattern as a data body has a definable positional relationship, the central position of each mesh on the mesh pattern on the image data can be obtained from the defined positional relationship and position data of the sampling reference pattern obtained on the image data. This will be described below with reference to FIGS. 20 to 22. In this case, a description will be made under an assumption that a scanning speed and a scanning direction of the image sensor with respect to an encoded image include certain variations. For the sake of simplicity, two synchronous mark arrays including clock marks and separated at a given interval are used as sampling reference patterns. A mesh pattern is arranged to be sandwiched between these two synchronous mark arrays.

FIG. 20 shows the positional relationship on image data read from the image sensor. Reference symbols  $U_i$  and  $U_{i+1}$  denote the central positions of  $i$ th and  $(i+1)$ th clocks of the upper synchronous mark array; and  $B_j$  and  $B_{j+1}$ , the central positions of  $j$ th and  $(j+1)$ th clocks of the lower synchronous mark array. For example, the central positions  $U_i$  and  $B_j$ , and  $U_{i+1}$  and  $B_{j+1}$  can be those of adjacent clocks which have a minimum sub-scanning distance on the read image data. In FIG. 20, the coordinates of the point  $U_i$  are represented by an origin (0,0), and the coordinates of the point  $U_{i+1}$  are represented by a point  $(x1,0)$  on the X-axis. The positions on the image memory are merely normalized and expressed for the sake of simplicity. The Y-axis direction corresponds to a direction of a main scanning line image.

FIG. 21 shows the positional relationship on a recording medium in correspondence with FIG. 20. Reference symbols  $T_i$  and  $T_{i+1}$  denote the centers of  $i$ th and  $(i+1)$ th clocks in the upper synchronous mark array;  $A_j$  and  $A_{j+1}$ , the centers of  $j$ th and  $(j+1)$ th clock marks in the lower synchronous mark array. The clock centers  $\{T_i\}$  and  $\{A_j\}$  of the synchronous mark arrays on the recording medium, and a central position  $\{P\}$  of each mesh of the reference pattern are already known items for the data reading apparatus, and the apparatus has these data as an encoded image model. For example, data of all these positions may be stored, or when intervals between positions have a predetermined synchronous relationship, all the central positions can be calculated from one or a plurality of position data and synchronous data. In FIG. 21, a line connecting the centers  $T_i$  and  $T_{i+1}$  on the upper synchronous mark array and a line connecting the centers  $A_j$  and  $A_{j+1}$  on the lower synchronous mark array are illustrated to be parallel to each other, but need not always be parallel to each other. In FIG. 21, the center of the  $i$ th clock of the upper synchronous mark array is represented by the origin (0,0), and the center of the  $(i+1)$ th clock is represented by a point  $(D,0)$  on the X-axis.

Assume that clock centers of the synchronous mark arrays as the reference patterns are obtained from the read image data by the method as described above, and a rectangular image block surrounded by the four clock centers  $U_i$ ,  $U_{i+1}$ ,  $B_j$ , and  $B_{j+1}$  (FIG. 20) is considered. The problem is how to detect the central positions of meshes present in this rectangular image block. Thus, the data reading apparatus selects corresponding clock centers  $T_i$ ,  $T_{i+1}$ ,  $A_j$ , and  $A_{j+1}$  on the encoded image model (i.e., on the recording medium). Since the encoded image model includes position data  $\{P\}$

of the centers of meshes in the entire encoded image, a set of central positions of meshes present in a rectangle (in this case, a parallelogram) defined by  $T_i, T_{i+1}, A_j$ , and  $A_{j+1}$  can be formed. For example, in the case of FIG. 21, a set of coordinates  $(x,y)$  which satisfy the following relation can be selected from a set  $\{P\}$  of central positions of all the meshes:

$$0 \leq x+Ly \leq D$$

( $L$  is the inclination)

Internal points of the rectangle defined by  $T_i, T_{i+1}, A_j$ , and  $A_{j+1}$  correspond to those in a rectangle defined by  $U_i, U_{i+1}, B_j$ , and  $B_{j+1}$  on the read image data, and never fall outside it. When the central positions of meshes present inside the rectangle are detected on the  $x$ - $y$  coordinate system of the encoded image model in this manner, the detected positions are coordinate-converted to positions on the  $X$ - $Y$  coordinate system of the read image data, and image bits at the converted position can be extracted.

FIG. 22(a)-22(f) show examples of coordinate conversion. FIG. 22(a) shows an image block of interest on the model, and the center of a certain mesh is represented by coordinates  $(x,y)$ . FIG. 22(f) shows a corresponding image block (target image block) on the read image data, and coordinates corresponding to the coordinates  $(x,y)$  are represented by  $(X,Y)$ . An operation executed between FIGS. 22(a) and 22(b) is horizontal scaling for matching the length of the upper side of a rectangle, an operation executed between FIGS. 22(b) and 22(c) is conversion for matching the inclination with that of the lower side (from  $U_i$  to  $B_j$ ) of the target image block, and an operation between FIGS. 22(c) and 22(d) is scaling for matching the length of the left side with that of the left side of the target image block. With these operations, of the four apexes  $U_i, U_{i+1}, B_j$ , and  $B_{j+1}$  of the target image block, three points  $U_i, U_{i+1}$ , and  $B_j$  coincide with corresponding points of the image block of interest. As shown in FIG. 22(e), a corresponding point of the image block shown in FIG. 22(d) need only be shifted to the point  $B_{j+1}$  diagonal to the point  $U_i$ , thus obtaining the target image block. An image block is a local area, and changes in speed or direction in the local area are very small even in manual scanning. Therefore, conversion from FIG. 22(d) to FIG. 22(e) can be considered as a uniform distortion (distortion proportional to an area ratio) without posing any problem.

As a result, the target coordinates  $(X,Y)$  are given by:

$$X = X_d + \Delta D \cdot \left( \frac{X_d + M \cdot Y_d}{X_1} \right) \cdot \left( \frac{Y_d}{Y_2} \right) \quad (1)$$

$$Y = Y_d + \Delta H \cdot \left( \frac{X_d + M \cdot Y_d}{X_1} \right) \cdot \left( \frac{Y_d}{Y_2} \right) \quad (2)$$

for

$$\Delta D = (X_3 - X_2)X_1 \quad (3)$$

$$\Delta H = Y_3 - Y_2 \quad (4)$$

$$X_d = \left( \frac{X_1}{D} \right) x + (M - L) y \quad (5)$$

$$Y_d = \left( \frac{Y_2}{H} \right) y \quad (6)$$

Multiplicators for  $\Delta D$  and  $\Delta H$ , given by the above equations (1) and (2) represent an area ratio of a rectangle in FIG. 22(d) indicated by a solid line, and a rectangle indicated by a dashed line and having coordinates  $(X_d, Y_d)$  as a diagonal point.  $\Delta D$  is the difference between the lengths of the upper and lower sides of the target image block, and  $\Delta H$  is the difference between the lengths along the  $Y$ -axis of the right and left sides of the target image block (see FIG. 20 for

further details about  $X_1, X_2, X_3, Y_1$ , and  $Y_2$ ). In this manner, the central coordinates  $(X,Y)$  of a corresponding mesh in the target image block can be obtained from the coordinates  $(x,y)$  of a mesh center in the model. Therefore, an image bit present at the position  $(X,Y)$  is sampled to obtain data indicating a black or white level of the mesh. Note that since an image frame formed by the points  $T_i, T_{i+1}, A_j$ , and  $A_{j+1}$  is actually an elongated small local area (see FIG. 2), an offset  $\Delta H$  is normally considered to have a negligible magnitude ( $\Delta H=0$ ) even when manual scanning is performed. Therefore, equation (2) can be simplified as  $Y=Y_d$ .

As described above, according to one arrangement of the present invention, image data obtained by reading an encoded image on a recording medium is segmented into a plurality of partial image data blocks, and a feature pixel group featuring a piece of a main scanning reference pattern is searched using the segmented partial image data blocks as searching units. A position of the reference scanning pattern in a partial image data block for which searching is unsuccessful is determined based on position data of normal pieces, indicated by the feature pixel groups for which searching is successful, of the main scanning reference pattern. Therefore, even when the main scanning reference pattern is partially destroyed by, e.g., contamination on the recording medium, its position can be obtained with high precision, and black and white levels of meshes of the mesh pattern can be reliably identified on the basis of the position data of the main scanning reference pattern.

Furthermore, according to one aspect of the present invention, when synchronous marks used for positioning a mesh pattern as elements of a scanning reference pattern are searched from image data obtained by reading an encoded image including the mesh pattern and the scanning reference pattern, a position of a synchronous mark to be determined next is predicted according to feature parameters of a synchronous mark determined by actual measurement, and the synchronous mark is actually measured to have the predicted position as the center. The actual measurement result and the predicted content are compared with each other. When the two results essentially coincide with each other, the actually measured position is determined as the position of the synchronous mark. However, when a non-coincidence is found, a mark destroyed portion caused by, e.g., contamination is detected, and the position of the synchronous mark is determined based on the already determined feature parameters. Therefore, even when the scanning reference pattern is partially destroyed by contamination or stain, the position of the corresponding synchronous mark can be determined with high precision without omission. Thus, black and white levels of meshes of the mesh pattern can be reliably identified based on the determined position data.

#### Third Embodiment

The third embodiment of the present invention will be described below. <Data Recording Apparatus>

FIG. 23 shows the overall arrangement of a data recording apparatus 1000 according to the third embodiment. The object of the entire apparatus is to record an image obtained by encoding data on a recording medium such as paper or a sheet (which image is the same as that shown in FIG. 2 described in the first embodiment, and will be described below with reference to FIG. 2). The characteristic feature of the data recording apparatus will be described in detail later.

In FIG. 23, a CPU 1010 is operated according to a program stored in a ROM 1020. In a data loading mode, the CPU 1010 loads target data (e.g., a user program list) supplied from an appropriate external device 1050 (e.g., a

23

personal computer) in a RAM 1030 through a communication interface 1040. In a recording mode, the CPU 1010 performs various error countermeasures of the target data stored in the RAM 1030 to form a data array which has a one-to-one correspondence with image elements (meshes in this case) of an encoded image. Thereafter, the CPU 1010 controls a printer 1060 to print the encoded image on a recording medium.

The operation of the data recording apparatus 1000, especially error countermeasures, will be described in detail below with reference to FIGS. 24 to 33.

In step 20-1 in FIG. 24, target data is loaded from the external device 1050 to the RAM 1030 through the communication interface 1040. The bit matrix of this target data is denoted by reference numeral 300 in FIG. 25. When the encoded image 20 shown in FIG. 2 described in the first embodiment is assumed to be used, the format of the mesh pattern 22 as its data body corresponds to a 48×72 two-dimensional mesh matrix. Each mesh is an image element expressing one bit as a minimum unit of data. For example, a black mesh expresses a bit "1", and a white mesh expresses a bit "0". When the 48×72 mesh pattern 22 is used, the size of the target data is given by 8 (bits)×256 (256 bytes), and the data recording apparatus 1000 processes 2,048-bit target data  $d_0$  to  $d_{2047}$  as the 64 (bits)×32 (bits) two-dimensional bit matrix 300, as shown in FIG. 25.

In step 20-2 (FIG. 24), 8 bits in the bit matrix 300 are considered as 1 symbol, and target data  $s_0$  to  $s_{255}$  consisting of 256 symbols are considered as an 8×32 two-dimensional symbol matrix 400 shown in FIG. 26. The size of each symbol is significant as a data unit (word) in steps 20-3 and 20-4 (to be described later). Note that the present invention is not limited to 1 symbol=8 bits in step 20-2.

The following relationship is established between the two-dimensional symbol matrix 400 and the two-dimensional bit matrix 300:

$$S_n = d_{8n+7} + 2d_{8n+6} + 2^2d_{8n+5} + 2^3d_{8n+4} + 2^4d_{8n+3} + 2^5d_{8n+2} + 2^6d_{8n+1} + 2^7d_{8n}$$

In step 20-3 (FIG. 24), error checking codes are added to the two-dimensional symbol matrix 400 of the target data to form a symbol matrix 500 with error checking codes shown in FIG. 27. Arbitrary codes may be used as error checking codes. In FIG. 27, four checking symbols  $p$  are added to each flow (line) of the symbol matrix 400 of the target data, and four checking symbols  $q$  are added to each of 12 columns of a two-dimensional matrix formed by the symbols  $p$ . In this case, since four error checking symbols (block checking codes) are added per block (one flow or column), the four checking codes themselves have a double symbol error correction function. Furthermore, since two systems of checking codes  $p$  (flows) and  $q$  (columns) are added to the symbols  $s$ , quadruple erasure correction in each block of the target data can be realized. In the matrix shown in FIG. 27, checking symbols  $p$  and  $q$  are added assuming so-called Reed-Solomon codes. Its relation is as follows:

24

$$\begin{bmatrix} \alpha^0 & \alpha^0 & \dots & \alpha^0 & \alpha^0 & \alpha^0 \\ \alpha^{11} & \alpha^{10} & \dots & \alpha^2 & \alpha^1 & \alpha^0 \\ \alpha^{22} & \alpha^{20} & \dots & \alpha^4 & \alpha^2 & \alpha^0 \\ \alpha^{33} & \alpha^{30} & \dots & \alpha^6 & \alpha^3 & \alpha^0 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \\ s_8 \\ s_9 \\ s_{10} \\ s_{11} \\ s_{12} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (7)$$

and,

$$\begin{bmatrix} \alpha^0 & \alpha^0 & \dots & \alpha^0 & \alpha^0 & \alpha^0 \\ \alpha^{35} & \alpha^{34} & \dots & \alpha^2 & \alpha^1 & \alpha^0 \\ \alpha^{70} & \alpha^{68} & \dots & \alpha^4 & \alpha^2 & \alpha^0 \\ \alpha^{105} & \alpha^{102} & \dots & \alpha^6 & \alpha^3 & \alpha^0 \end{bmatrix} \begin{bmatrix} s_{16} \\ s_{17} \\ s_{18} \\ s_{19} \\ s_{20} \\ s_{21} \\ s_{22} \\ s_{23} \\ s_{24} \\ s_{25} \\ s_{26} \\ s_{27} \\ s_{28} \\ s_{29} \\ s_{30} \\ s_{31} \\ s_{32} \\ s_{33} \\ s_{34} \\ s_{35} \\ s_{36} \\ s_{37} \\ s_{38} \\ s_{39} \\ s_{40} \\ s_{41} \\ s_{42} \\ s_{43} \\ s_{44} \\ s_{45} \\ s_{46} \\ s_{47} \\ s_{48} \\ s_{49} \\ s_{50} \\ s_{51} \\ s_{52} \\ s_{53} \\ s_{54} \\ s_{55} \\ s_{56} \\ s_{57} \\ s_{58} \\ s_{59} \\ s_{60} \\ s_{61} \\ s_{62} \\ s_{63} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (8)$$

where  $\alpha$  as a primitive element of a Galois body  $GF(2^8)$  is one of roots of a polynomial given by:

$$x^8 + x^4 + x^2 + x + 1$$

$\alpha$  takes a value of 2 in decimal notation, and is expressed by a binary number as follows:

$$00000010(B)$$

In this connection,  $\alpha^8 = \alpha^4 + \alpha^3 + \alpha^2 + 1 = 16 + 8 + 4 + 1 = 29$ , and  $\alpha^1$  to  $\alpha^{255}$  take corresponding proper values from 1 to 255 ( $\alpha^{255} = 1$ ). "0"s of the right-hand sides of equations (7) and (8) indicate that a so-called syndrome is zero.

For example, when equation (7) is solved for  $q$ , we have:

$$[Q] = [A_1]^{-1} \cdot [A_2] \cdot [W] \quad (9)$$

for

$$[Q] = \begin{bmatrix} q^0 \\ q^1 \\ q^2 \\ q^3 \end{bmatrix}$$

$$[A_2] = \begin{bmatrix} \alpha^0 & \alpha^0 & \dots & \alpha^0 & \alpha^0 & \alpha^0 \\ \alpha^{11} & \alpha^{10} & \dots & \alpha^5 & \alpha^4 & \alpha^3 \\ \alpha^{22} & \alpha^{20} & \dots & \alpha^{12} & \alpha^{10} & \alpha^8 \\ \alpha^{33} & \alpha^{30} & \dots & \alpha^{18} & \alpha^{15} & \alpha^{12} \end{bmatrix}$$

25

-continued

$$[A_1]^{-1} = \begin{bmatrix} \alpha^0 & \alpha^0 & \alpha^0 & \alpha^0 \\ \alpha^3 & \alpha^2 & \alpha^1 & \alpha^0 \\ \alpha^6 & \alpha^4 & \alpha^2 & \alpha^0 \\ \alpha^9 & \alpha^6 & \alpha^3 & \alpha^0 \end{bmatrix}^{-1} = \begin{bmatrix} \alpha^{212} & \alpha^{153} & \alpha^{152} & \alpha^{209} \\ \alpha^{156} & \alpha^2 & \alpha^{125} & \alpha^{152} \\ \alpha^{158} & \alpha^{138} & \alpha^2 & \alpha^{153} \\ \alpha^{218} & \alpha^{158} & \alpha^{156} & \alpha^{212} \end{bmatrix}$$

$$[W] = \begin{bmatrix} s_0 \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \\ s_7 \end{bmatrix}$$

Equation (8) is similarly solved to calculate p.

As a result of adding the error checking codes p and q, the two-dimensional matrix 500 having a size of 12×36 symbols can be obtained.

The above-mentioned block type error checking codes p and q are mainly used as countermeasures against random errors, and are not effective for burst errors. For example, when a contamination becomes attached to a considerably wide contiguous range on an encoded image, an error beyond correction capability of the error checking codes may occur.

In step 20-4 in FIG. 24, countermeasure processing against burst errors is executed to change the order of elements of the matrix 500 so that a contiguous area on the two-dimensional symbol matrix 500 obtained in step 20-3 (processing for adding error checking codes) is converted into discrete small areas on the encoded image. Processing in step 20-4 will be referred to as scrambling or interleaving hereinafter. FIGS. 28A, 29A, and 30A respectively show symbol matrices 600, 700, and 800 after first, second, and third scrambling operations. Symbol matrices 600S, 700S, and 800S in FIGS. 28B, 29B, and 30B are simplified ones of the structures of the symbol matrices 600, 700, and 800 to allow easy understanding. In the first scrambling operation, line data of the symbol matrix 500 after the error checking codes are added is replaced. That is, ninth and tenth line data of the original matrix are moved to the third and fourth lines of the original matrix 500, respectively; the third to sixth line data of the original matrix 500 are moved downward by two lines; eleventh and twelfth line data of the original matrix 500 are moved to the ninth and tenth lines of the original matrix 500, respectively; and seventh and ninth line data of the original matrix 500 are moved to the eleventh and twelfth lines, respectively, thereby forming a first scrambled matrix 600. Please refer to the matrix 600S in FIG. 28B as a rough structure of the matrix 600 shown in FIG. 28A. In the second scrambling operation after the first scrambling operation, the lower half of the first scrambled matrix 600 is attached to the right-hand side of the upper half, thus forming a 6×72 (symbols) second scrambled matrix 700. Finally, in the third scrambling operation, the second scrambled matrix 700 is moved in the horizontal direction by 16 symbols. As a result, a third scrambled matrix 800 shown in FIG. 30A is obtained.

The last scrambled matrix 800 has the same positional relationship as the two-dimensional mesh matrix on the encoded image, and contiguous elements on the scrambled matrix 800 are similarly contiguous elements on the encoded image. However, the matrix 800 has the relationship with the

26

matrix 500 before scrambling so that a contiguous portion on the matrix 500 is located at discrete positions on the scrambled matrix 800. For example, for  $q_{128}$ ,  $q_{129}$ ,  $q_{130}$ , and  $q_{131}$  in the ninth to twelfth lines of the 33rd column on the matrix 500 before scrambling,  $q_{128}$  and  $q_{129}$  are located at the third and fourth lines of the 17th column, and  $q_{130}$  and  $q_{131}$  are located at the third and fourth lines of the 53rd column on the scrambled matrix 800. As a result, since a burst error on the encoded image is scattered as short errors on the matrix 500 before scrambling, error correction within the capability of the error checking codes p and q can be realized. As schematically shown by the matrix 800S in FIG. 30B, a q portion occupies the central portion of the scrambled matrix 800 (FIG. 30A). This arrangement is made in consideration of a method of adding the error checking codes (FIG. 27) and a mesh data sampling method of a data reproducing apparatus. More specifically, as can be seen from FIG. 27, two systems of error checking codes p and q are associated with target data s, while only codes for checking horizontal blocks (lines) are added to the error checking codes q. For this reason, even if a code error occurs in this portion, there is no means for confirming the position of the error. As long as block checking codes or finite convolution checking codes are used on a finite matrix, this problem is not perfectly avoided. In other words, on the matrix 500 shown in FIG. 27, an error at a symbol  $s_0$  reflected on both a syndrome value of the first line and a syndrome value of the first column (symbol values are not equal to zero), while an error at a symbol  $q_{128}$  is reflected on only a syndrome value of a column (33rd column) to which  $q_{128}$  belongs. For this reason, when errors occur at three positions, e.g., the symbols  $q_{128}$ ,  $q_{129}$ , and  $q_{130}$ , correction is impossible. An error in an area of q is serious, and hence, its frequency is preferably decreased to be lower than that in other portions. Meanwhile, in main scanning decoding processing described later in a data reproducing apparatus, image bits at points obtained by equally dividing an interval between the guide lines 21 are sampled using the upper and lower guide lines 21 of the mesh pattern 22 (FIG. 2) as the main scanning reference patterns. In this sampling method, an offset of an equally divided point from a true position is maximized near each guide line 21, and the central portion of the mesh pattern 22 has maximum safety. Therefore, as shown in FIG. 30A, the error checking codes q are arranged at the central portion of the matrix 800, thus decreasing an error rate of these codes q to be lower than those of other codes.

The scrambled two-dimensional symbol matrix 800 is converted to a 48×72 (bit) two-dimensional bit matrix 900 while considering each symbol as vertical 8 bits, as shown in FIG. 31 (step 20-5 in FIG. 24). In this case, for example,  $s_{128} = [b_0 b_1 b_2 b_3 b_4 b_5 b_6 b_7]^T$  (where T is the transposed matrix).

Step 20-6 is executed to prevent an error caused by a variation of conversion characteristics of the image sensor used in a data reproducing apparatus. More specifically, photoelectric conversion characteristics of an image sensor cannot perfectly follow an instantaneous value of incident light. Thus, when an image is scanned using such an image sensor, a conversion output varies between a case wherein scanning enters a white (or black) area after a long black (or white) area, and a case wherein a white (or black) area is kept scanned. For this reason, a white portion isolated in a black area is affected by the surrounding black area, and is detected while being reduced, or may sometimes be detected as a black pixel. As a result, accurate image data cannot often be provided. As shown in FIG. 2, the encoded image of the present invention comprises a two-dimensional black-

and-white mesh matrix of the mesh pattern 22, and each black or white mesh expresses one bit. Therefore, it is preferable that pixel values are accurately obtained, and this is necessary when each mesh is small and hence, a recording density is high. Meanwhile, since data often have the same bit value, the above-mentioned detection error of a pixel value tends to occur in such an area. In step 20-6 in FIG. 24, bit values on the scrambled two-dimensional bit matrix 900 are randomized using pseudo random numbers, so that the same bit value does not continue over a long distance, thereby solving the above problem.

FIG. 32 shows an example of randomization. A pseudo random number generator P-RND EXORs an output from a D flip-flop D15 of a 16-bit shift register (constituted by D flip-flops 1 to 16) and an output from a D flip-flop D1, shifts bits of the shift register to the right, and inputs the EXORed result to the D flip-flop D1. A pseudo random number output  $rnd(n)$  is extracted from the D flip-flop D16, and is EXORed with a corresponding element  $b_n$  on the bit matrix 900 ( $\bar{b}_n = b_n \oplus rnd(n)$ ), thereby randomizing the elements of the matrix 900. At the beginning of the operation, an appropriate initial value (e.g., BB8 (hex)) is preferably set in the shift register. As a result, a  $48 \times 72$  randomized two-dimensional bit matrix 1100 is formed, as shown in FIG. 33.

The randomized two-dimensional bit matrix 1100 has a one-bit to one-mesh correspondence with the two-dimensional mesh matrix of the mesh pattern 22 to be recorded on the recording medium while positions on the matrices coincide with each other, and black and white levels of the meshes are specified by bit values.

In order to achieve this, in step 20-7, the mesh pattern 22 according to the bit matrix 1100 is printed on the recording medium. In step 20-7, sampling reference marks having a predetermined positional relationship with the mesh pattern 22 are also printed. The sampling reference marks adopt the guide lines 21, the synchronous mark arrays 22, and the data start and end marks 23 and 24 shown in FIG. 2. As will be described later, a data reproducing apparatus finds these sampling reference marks from read data of the encoded image 20, determines mesh positions in the mesh pattern 22 with reference to positions of these reference marks, and samples image bits there. <Data Reproducing Apparatus>

A method and apparatus for reading the encoded image 20 (FIG. 2) recorded on a recording medium described above have already been described in association with the first embodiment, and a repetitive description will be avoided.

This embodiment will be described below with reference to the circuit arrangement shown in FIG. 1. A CPU 13 operates according to a program stored in a ROM 14 (see FIG. 34). While an image sensor 11 scans an image, every time a controller 12 converts serial image data into parallel data, the CPU 13 writes the parallel data in an image RAM 15 (14-1). Upon completion of image scanning, the CPU 13 starts a decoding operation (14-2 to 14-6) of image data stored in the image RAM 15, and stores the decoded result in a RAM 16.

The decoding operation for reproducing data is basically achieved by following encoding processing executed by the data recording apparatus 1000 in an opposite direction. First, main scanning sampling step 14-2 (FIG. 34) and sub-scanning sampling step 14-3 are executed to sample image bits representing black and white levels of meshes located at the central positions of meshes from read image data, thereby obtaining a matrix corresponding to the last two-dimensional bit matrix 1100 described in data recording. Processing operations in steps 14-1 to 14-3 are the same as those described in association with the first embodiment. In

DC (derandomization) step 14-4, the elements on the matrix are derandomized using the pseudo random numbers described with reference to FIG. 32, thereby forming a matrix corresponding to the non-randomized bit matrix 900 shown in FIG. 31. The reason why the non-randomized bit matrix 900 is restored is that the randomization is given by:

$$\bar{b}_n = b_n \oplus rnd(n)$$

derandomization is given by:

$$b_n = \bar{b}_n \oplus rnd(n)$$

and, the two equations are derived for the same pseudo random number  $rnd(n)$ . The DC processing itself is the same as randomization processing step 20-6, except that data to be EXORed with  $rnd(n)$  is  $b_n$  (non-randomized data) or  $\bar{b}_n$  (randomized data), and a further description thereof will be omitted.

After the non-randomized symbol matrix is obtained, processing (descrambling) Opposite to scrambling processing step 20-4 described with reference to FIGS. 28A and 28B, FIGS. 29A and 29B, and FIGS. 30A and 30B is executed to descramble the symbols, thereby obtaining a matrix corresponding to the symbol matrix 500 with the error checking codes before scrambling (14-5). This descrambling processing step 14-5 can also be understood from the description about the scrambling processing, and a detailed description thereof will be omitted.

A difference between the descrambled symbol matrix with the error checking codes and the symbol matrix with the error checking codes which is formed in encoding is an error. Thus, in step 14-6, error correction is executed using the error checking codes, thus obtaining target data.

Of the series of these processing operations, a matrix, in which black and white levels of meshes are indicated by bits, corresponding to the last two-dimensional bit matrix 1100 is obtained as a result of sampling based on the scanning reference patterns. Under a user environment wherein a scanning speed and a scanning direction can be kept constant using a mechanical scanning type image sensor, even if no scanning reference patterns are arranged, meshes can be directly recognized from image data which is free from image distortion, as has been described in the second embodiment.

The error correction step 14-6 as the characteristic feature of this embodiment will be described in detail below. FIG. 36 shows the entire error correction flow. Before this flow starts, a descrambled two-dimensional symbol matrix 500R with error checking codes is obtained, as shown in FIG. 35. P0 to P7 in FIG. 35 respectively indicate first to eighth lines of the two-dimensional matrix 500R. Since error checking codes  $p$  are added to these lines, these lines are called P vectors in FIG. 36. Since error checking codes  $q$  are added to columns (indicated by Q0 to Q35), of the two-dimensional matrix, these columns are called Q vectors in FIG. 36. In the loop including steps 24-1 to 24-3 in FIG. 36, error correction processing step 24-2 is executed once for the 36 Q vectors. Thereafter, in the loop including steps 24-4 to 24-6, error correction processing step 24-5 is executed once for the eight P vectors. Thereafter, the flow advances to step 24-7 to check if correction is actually executed in steps 24-2

and 24-5 or erasure registration necessary for subsequent correction is performed although no correction is executed. Since it can be considered that an error remains excluding a case wherein neither of these operations are performed, a correction count is reset to zero, and steps 24-1 to 24-5 are

executed under a condition that the number of loop times is small (5 or less in FIG. 36). When the number of loop times is large, considerable errors are included in the matrix 500R, and even if error correction is continued, error correction must be further repeated. Thus, the processing is stopped in this case.

FIG. 37 shows error correction processing steps 24-2 and 24-5 in detail. In step 25-1, calculations corresponding to the left-hand side of equation (7) are executed for the Q vectors, and calculations corresponding to the left-hand side of equation (8) are executed for the P vectors, thereby obtaining syndrome values. That is, a syndrome matrix is multiplied with a vector of interest. If there is no error in the vector (block) of interest, syndrome values ( $S_0, S_1, S_2, S_3$ ) become zero, as shown in equations (7) and (8). Thus, the syndrome values are checked in step 25-2. If all the syndrome values are zero, the flow returns to the main routine since no error is detected. Thus, an operation of checking a no-error state which has the highest possibility is completed. An error having the second highest possibility is a single-symbol error, and an error having the third highest possibility is double-symbol errors. Since four error checking codes are added per vector, correction can be made up to the double errors. When the syndrome values become non-zero, an error of one or two symbols is assumed. If two symbols in one vector include errors, location values representing error positions are represented by  $X_1$  and  $X_2$ , and error patterns are presented by  $Y_1$  and  $Y_2$ , the following relations are established for the syndrome values  $S_0, S_1, S_2$ , and  $S_3$  which have already been calculated in step 25-1:

$$S_0 = Y_1 + Y_2$$

$$S_1 = X_1 Y_1 + X_2 Y_2$$

$$S_2 = X_1^2 Y_1 + X_2^2 Y_2$$

$$S_3 = X_1^3 Y_1 + X_2^3 Y_2$$

Thus, four simultaneous equations for four unknowns  $X_1, X_2, Y_1$ , and  $Y_2$  are obtained. The location values  $X_1$  and  $X_2$  become a root of  $AX^2+BX+C=0$  for:

$$A = S_0 S_2 + S_1^2$$

$$B = S_1 S_2 + S_0 S_3$$

$$C = S_1 S_3 + S_2^2$$

In step 25-3, values of A, B, and C are calculated. In consideration of a possibility of a single-symbol error, since  $Y_1$  or  $Y_2$  is zero,  $A=0$  is established. Thus, it is checked in step 25-4 if  $A=0$ . If  $A=0$ , a single-symbol error is determined, and its error pattern ( $Y=S_0$ ) and a location value ( $X=S_1/S_0$ ) are obtained. The error pattern is added (EXORed) to a symbol specified by the location value, thus correcting the single-symbol error. Thereafter, a correction count is incremented by one (25-5 to 25-7).

If  $A \neq 0$  is not established,  $D (=B/A)$  and  $E (=C/A)$  are calculated (25-8), and the following quadratic expression is solved (25-9):

$$(X/D)^2 + (X/D) = E/D^2 \quad (10)$$

If double-symbol errors occur, two real roots are obtained. If the two real roots are obtained in step 25-10, double-symbol errors are determined, and error location values  $X_1$  and  $X_2$  ( $X_1$  is D times of a solution R of equation (10), i.e.,  $DR$ , and  $X_2$  is given by  $X_2 = D/X_1$ ), and error patterns  $Y_1$  and  $Y_2$  ( $Y_2 = (X_1 S_0 + S_1)/(X_1 + X_2)$ ,  $Y_1 = S_0 + Y_2$ ) are calculated. The

error patterns are respectively added to the two symbols indicated by the two error location values to correct these symbols, and the correction count is incremented by two (25-11 to 25-13). If NO is determined in step 25-10, since three or more errors occur and cannot be corrected by this correction alone, the flow advances to erasure processing step 25-14 utilizing results of vectors in other systems.

FIG. 38 shows an example of erasure processing step 25-14. In step 26-1, erasure-registered positions of a vector of interest are counted. For example, if vectors  $P_0$  to  $P_2$  are erasure-registered for the vector  $Q_0$  in FIG. 35, the positions of the symbols  $S_0, S_1$ , and  $S_2$  have a high possibility of an error. It is checked in step 26-2 if the number of erasure-registered position is 3. If there is no erroneous correction so far, symbols at positions which are not erasure-registered must be correct. Therefore, if triple-symbol errors occur, their positions must be erasure-registered. Thus, if YES is determined in step 26-2, triple-symbol errors are determined, and error patterns are calculated using the three erasure-registered positions as error locations. For example, if the three error location values are represented by  $X_1, X_2$ , and  $X_3$ , and their error patterns are represented by  $Y_1, Y_2$ , and  $Y_3$ , the following equations are established:

$$S_0 = Y_1 + Y_2 + Y_3$$

$$S_1 = X_1 Y_1 + X_2 Y_2 + X_3 Y_3$$

$$S_2 = X_1^2 Y_1 + X_2^2 Y_2 + X_3^2 Y_3$$

Since only the error patterns  $Y_1, Y_2$ , and  $Y_3$  are unknown in these three simultaneous equations, their solutions can be obtained. For example,  $Y_3$  can be calculated by:

$$Y_3 = \frac{X_1 X_2 S_0 + (X_1 + X_2) S_1 + S_2}{(X_1 + X_2)(X_2 + X_3)}$$

In step 26-5, error patterns are respectively added to three symbols on the matrix associated with errors to correct these symbols, and the correction count is incremented by 3 in step 26-6.

If the number of erasure-registered positions is not 3 in step 26-2, the position of the vector of interest is erasure-registered to give an opportunity of erasure processing by vectors of another system (Q vector) different from the vector of interest (e.g., P vector) (26-7).

FIGS. 39(a)-39(g) show examples of burst errors on the encoded image 20. In FIGS. 39(a) to 39(g), contiguous black areas indicate burst errors. According to this embodiment, target data could be correctly obtained even if any burst errors shown in FIGS. 39(a) to 39(g) occur.

Embodiments of the present invention have been described. Various other changes and modifications may be made within the spirit and scope of the invention. For example, the data recording apparatus and the data reproducing apparatus can be easily integrated to constitute a recording/reproduction unit. The mesh pattern 22 as a two-dimensional mesh matrix without any gap is used as a data body of the encoded image 20 but may be a two-dimensional matrix of image elements including gaps. A photoelectric conversion type image sensor is employed in the above embodiments. However, a magnetic type image sensor which responds to a magnetic ink or the like may be used. In the above embodiments, scrambling (interleaving) processing and randomization processing of element values are separately executed. For example, position scrambling for moving element positions (ij) on a two-dimensional matrix to other positions using a random number table may be executed to simultaneously remove DC components.

According to one arrangement of the present invention, when an encoded image on a recording medium is constituted by a two-dimensional matrix of image elements expressing data elements, error checking codes are added to a two-dimensional data matrix constituting target data as original information, and the encoded image as the two-dimensional matrix of image elements is printed on the recording medium according to the two-dimensional data matrix with the error checking codes. Therefore, a recording medium which is not easily influenced by errors can be provided even at a relatively high recording density. In addition, printing precision can be reduced according to error correction capability of error checking codes.

Furthermore, according to one arrangement of the present invention, data of an encoded image is read from a recording medium, image elements expressing data elements are detected from the read image data to obtain a two-dimensional data matrix, and an error on the matrix is corrected according to error checking codes included in this matrix, thereby reproducing target data. Thus, target data can be correctly reproduced even if an error occurs on a two-dimensional data matrix when an encoded image is read, when image elements are recognized from the image data, or due to error factors before these processes.

Since a means for changing the order of a data matrix with error checking codes in recording and for reversing the order to restore the original data matrix with the error checking codes in reproduction is arranged, characteristics strong against burst errors can be obtained. More specifically, even when errors continuously occur in recognition of image elements during reproduction due to a considerably large contamination on an encoded image on a recording medium, the burst errors can be converted to discrete errors on the original data matrix with the error checking codes. Therefore, these errors can be easily corrected within the error correction capability of the error checking codes. As a result, reproducibility of correct target data can be improved.

Furthermore, according to one arrangement of the present invention, a read error of image elements of an encoded image caused by a change in conversion characteristics of an image sensor, especially, a change in characteristics caused by DC components, i.e., continuation of the same density is taken into consideration. When an encoded image is recorded, element values of a data matrix constituting given data are randomized using pseudo random numbers. Upon reproduction, after the randomized data matrix is obtained, the data matrix is derandomized to restore element values of the original data matrix. Therefore, no long mesh arrays having the same density appear in the encoded image on the recording medium. When the encoded image is read, pixels can be accurately digitized (e.g., binarized), and an error in reproduced data caused by a read error can be prevented.

Some embodiments of the present invention have been described. However, these embodiments are merely examples, and the present invention may have various other arrangements. All the modifications and applications of the present invention are incorporated in the scope of the

invention, and, hence, the scope of the present invention should be determined by only the appended claims and their equivalents.

What is claimed is:

1. A data reproduction method comprising the steps of:

reading an encoded image as image data from a recording medium on which image elements including a mesh pattern having a rectangular shape as a whole and including dark and light image elements which are arranged at two-dimensional distributed positions, whose arrangement is determined based on a changed two-dimensional matrix obtained by changing an original two-dimensional matrix constituting target data in accordance with predetermined pseudo random numbers, respective randomized target data  $b_n$  being obtained by executing an exclusive OR operation for target data  $b_n$  with pseudo random numbers  $\text{rnd}(n)$ ;

detecting image elements from the read image data, and obtaining the elements of the randomized data to form the changed two-dimensional matrix; and

changing the elements on the formed two-dimensional matrix according to the predetermined pseudo random numbers to form the original two-dimensional matrix, thereby reproducing the target data, said target data  $b_n$  being obtained by executing an exclusive OR operation for the randomized target data  $b_n$  with the same pseudo random numbers  $\text{rnd}(n)$ .

2. A data reproduction apparatus comprising:

image sensor means for reading an encoded image as image data from a recording medium on which image elements including a mesh pattern having a rectangular shape as a whole and including dark and light image elements which are arranged at two-dimensional distributed positions, and whose arrangement is determined based on a changed two-dimensional matrix obtained by changing an original two-dimensional matrix constituting target data in accordance with predetermined pseudo random numbers, respective randomized target data  $b_n$  being obtained by executing an exclusive OR operation for target data  $b_n$  with pseudo random numbers  $\text{rnd}(n)$ ;

image element decoding means for detecting image elements from the read image data, and obtaining the elements from the read image data, and obtaining the elements of the randomized data to form the changed two-dimensional matrix; and

derandomizing means for changing the elements on the formed two-dimensional matrix according to the predetermined pseudo random numbers to form the original two-dimensional matrix, thereby reproducing the target data, said target data  $b_n$  being obtained by executing an exclusive OR operation for the randomized target data  $b_n$  with the same pseudo random numbers  $\text{rnd}(n)$ .

\* \* \* \* \*